Use your kind words please: Effects of tootling on student and teacher behaviour in New Zealand primary-school classrooms

A thesis

submitted in fulfilment

of the requirements for the degree

of

Master of Applied Psychology

(Behaviour Analysis)

at

The University of Waikato

by

Cheniel Powell

THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

2020

Abstract

The aim of this study was to extend the literature on *tootling*, an intervention consisting of student reports of peers' prosocial behaviour. Effects on student on-task behaviour and disruptions, alongside teacher praise statements, in a New Zealand primary-school were investigated. A single-subject multiple baseline design, with a follow-up phase, was used in an evaluation of the effects of the intervention across three groups. Students were taught how to notice and record peers' prosocial behaviour on slips of paper, which were collected each day before the teacher read several out loud to the class and provided feedback. An interdependent group contingency, whereby students worked together towards a shared goal, alongside public display of progress, was used to provide class-wide rewards for reaching a pre-determined tootle target. An increase in on-task behaviour across all three groups, decreases in student disruptions of one group, and little, to no change in teacher praise statements were found. Results mostly remained consistent during follow-up observations. Teachers and students rated the intervention socially acceptable and, according to treatment integrity scores, tootling can be used effectively in a classroom setting. Implications and future directions are discussed.

Acknowledgements

Thank you to my supervisor, Associate Professor Angelika Anderson, for your knowledge, guidance, and time throughout this journey. Your encouraging words kept me focussed and on-task, whilst our shared interest in the field kept me passionate and motivated. Thank you, Dr Rebecca Sargisson, for believing in me from day 1, encouraging me to keep moving forwards when I wanted to side-step, and for being available, even when you were thousands of kilometres away.

A special thank you to my participants and the participating school, for welcoming me into your classrooms and making observations fun. You made my thesis possible and I will always be grateful for that. To Nikau, thank you for reminding me why I'm doing this. Isaac, thank you for your assistance and being calm when I felt anything but calm. To Tina and the silver linings, understanding, and shared determination to succeed. It's been an interesting journey and one which I'm thankful to have shared with you.

Finally, to Molly, for your time, support, and for being the best nana my boys could ask for, thank you. To my husband, Michael, and our three sons, Samuel, Noah, and Joshua, thank you for your endless support, understanding, and encouragement. We navigated this journey together and I will be forever grateful for your everlasting patience and belief in me.

Table of Contents

List of Figures

List of Tables

Literature Review

Introduction

Behaviour management is essential in every classroom, affecting both students and teachers (Johansen et al., 2011; Reinke et al., 2013). Inappropriate student behaviour reduces time available for instruction (Lassen et al., 2006; Riley et al., 2011) and low-level problem behaviour is a major contributor to teacher exhaustion and burnout (Aloe et al., 2014; Reinke et al., 2013). Links have been found between children with increased levels of disruptive classroom behaviour and mental health difficulties, drug abuse, unemployment, criminal activity, homelessness, and poor relationships during adolescence and adulthood (Parsonage et al., 2014). When classroom behaviour is well managed, a positive classroom environment is maintained (Sieberer-Nagler, 2016) and more time is available for academic instruction (Reinke et al., 2013). Although New Zealand teachers recognise the importance of behaviour management, Johansen et al. (2011) asked if they believed they had adequate formal training in this area, and found that 83.8% felt they were insufficiently trained.

Previously preferred punitive, reactive, and exclusionary classroom management techniques used in New Zealand, such as office referrals, detentions, suspensions, and expulsions, which focus on problem behaviour, are unsuccessful (Elder & Prochnow, 2016). These techniques often result in increased problem behaviour, poor teacher-student and student-student relationships, and emotional exhaustion in teachers (Cooper & Jacobs, 2011; Reinke et al., 2013). Teachers who use a punitive and coercive approach to behaviour management, inadvertently promote student coercive behaviour (Cooper & Jacobs, 2011). Similarly, whilst inappropriate behaviour of some students will decrease because of punishment strategies, other students learn to decrease the probability of punishment delivery by not “getting caught” (Skinner et al., 2000). Additionally, punitive techniques do not teach replacement behaviours, meaning students may never learn appropriate alternatives (Cooper

et al., 2014). Consequently, learning environments suffer and levels of academic achievement decrease (Reinke et al., 2014).

Positive Behaviour for Learning: School-Wide

Alternatively, positive teacher-student interactions enhance learning environments by strengthening relationships and increasing levels of appropriate behaviour and academic performance (Bost & Riccomini, 2006; Gage et al., 2015). Educational policies encourage evidence-based practices for all areas of educational instruction, including classroom management (Plavnick & Ferreri, 2013). New Zealand's Positive Behaviour for Learning: School-Wide (PB4LSW) is based on the empirically supported Positive Behavioural Interventions and Support framework (see Sugai et al., 2016, for more) widely and successfully used in the United States since the 1990s (Ministry of Education, 2019). A key feature of the PB4LSW framework is its focus on positive behaviour, achieved through teaching appropriate skills, instead of punishing inappropriate behaviour (Savage et al., 2011). With its foundations in Applied Behaviour Analysis (ABA), PB4LSW uses a preventative approach to problem behaviour by changing the environment, systems, and practices within the school (Elder & Prochnow, 2016; Lane et al., 2013; Savage et al., 2011).

The central theme of ABA examines the function of behaviour to teach socially significant behaviour to replace inappropriate behaviour, thus eliminating the need for punitive techniques (Skinner, 1968). For example, a student who is reprimanded for calling out in class may not have learned how to raise their hand and wait to be called upon by the teacher. A reprimand may stop the behaviour temporarily but will not teach the student an alternate response: ABA addresses this issue. In ABA, new behaviours are taught through principles of operant conditioning, using reinforcement, whereby an appetitive consequence, delivered contingent upon a specific response increases the likelihood of that response occurring again (Skinner, 1968). For example, after explaining to the child in the example

above, the correct way to get the teacher's attention, the teacher may respond with praise, allowing the student to speak, after the student raises their hand and waits. This praise will likely reinforce hand raising, making it more likely to occur again in future. This focus on teaching positive behaviour shifts attention from reacting to misbehaviour, to looking for opportunities to reinforce appropriate behaviour (Lane et al., 2013). Thus, the principle of behaviour as a product of the environment is adopted (Elder & Prochnow, 2016; Parsonson, 2012). The use of ABA in classroom behaviour management contributes to a positive learning environment, reducing disruptive behaviour and reinforcing appropriate behavioural and academic learning (Parsonson, 2012).

PB4LSW is built around a hierarchical framework with primary, secondary, and tertiary level interventions (Lane et al., 2013; Savage et al., 2011). The primary tier encompasses all students, with all behavioural expectations taught and reinforced through school developed initiatives (Ministry of Education, 2019). Students identified as unresponsive to these primary tier initiatives receive secondary and tertiary tier supports (Lane et al., 2013). Secondary tier supports are targeted at approximately 15% of students who need additional support, teaching social skills and targeted behaviours, typically via small group interventions. The tertiary tier targets the minority (5%) of students who need intensive support for severe challenging behaviour, focussing on individual interventions specific to those students' needs (Ministry of Education, 2019). This dynamic system offers fluid and temporary supports during the second and third tiers which can be removed once an individual is equipped with the necessary tools to return to the primary tier (Lane et al., 2013). To experience success in this framework, teachers need evidence-based behaviour strategies, which are practical and effective (Chaffee et al., 2017). ABA uses empirically researched, function-based interventions, which rely on functional behaviour assessments to collect information on the antecedents (stimulus or setting before behaviour occurs), and

consequences of behaviour, to help determine when inappropriate behaviour is most likely to occur and how to prevent it, by reinforcing alternate behaviour (Fettig & Barton, 2014).

Evidence-based classroom strategies

An evidence-based strategy often used in classroom management is the token economy, whereby students earn tokens for appropriate behaviour which can be traded for desirable items (Niesyn, 2009; Soares et al., 2016). Although not function based, token economies focus on positive behaviour and have been used successfully; to improve undergraduate student participation (Boniecki & Moore, 2003); as part of a class-wide approach for decreasing inappropriate behaviour of pre-school-aged children (Filcheck et al., 2004); and in conjunction with posted classroom rules, response cost, and mystery motivators, to improve levels of disruptive behaviours of three male, second-grade students identified with Oppositional Defiance Disorder (Mottram et al., 2002). Soares et al. (2016) conducted a meta-analysis on the effect size of token economy use in contemporary classrooms and found medium to large effects when used to improve academic and behavioural outcomes. However, in a systematic evaluation of the use of token economies for improving classroom behaviour, it was found that the literature fails to provide sufficient evidence that token economies are best practice for classroom management (Maggin et al., 2011). After reviewing 24 studies, Maggin et al. (2011) argued that there was a lack of methodological rigour, along with insufficient reporting on participants and settings to allow replication. Some educators are also reluctant to use token economies because of the time needed to implement them and their need for consistency in implementation (Niesyn, 2009). Adding to the demand already placed on teachers could cause them to feel overwhelmed and burnt-out.

Peer-Mediated Interventions

The focus on prevention in the PB4LSW framework means effective, evidence-based strategies are utilised to support student learning without overwhelming teachers with demand (Lane et al., 2013; Reinke et al., 2013). Multiple competing stimuli in the classroom, such as instruction and large group monitoring, means teachers cannot attend to all children all of the time (Radley et al., 2019). Alongside prevention, peer-mediated interventions, whereby students are trained to implement behaviour change procedures aimed at their peers, reduce teacher demand (Harjusola-Webb et al., 2012; Kaya et al., 2015; Kohler & Strain, 1990). Kaya et al. (2015) argued that trained peer mediators may function as discriminative stimuli, indicating to students that reinforcement is available for appropriate responses. In other words, the presence of a peer mediator may prompt a student to respond in a way which has resulted in reinforcement, previously. A meta-analysis on peer-mediated interventions found them to be effective in teaching a variety of skills to various age-groups, in multiple settings (Kohler & Strain, 1990). Examples include; adolescents improving conversation skills of peers with autism spectrum disorder (Bambara et al., 2016); primary-school-aged students using differential reinforcement of other behaviour with extinction to improve on-task behaviour of peers with attention deficit hyperactivity disorder (Grauvogel-Macaleese & Wallace, 2010); and 5- to 6-year old students improving beginning reading skills of peers with emotional and behavioural difficulties, through peer-tutoring (Grauvogel-Macaleese & Wallace, 2010).

Cooperative learning

Cooperative learning (CL) is an example of a well-researched, peer-mediated intervention (PMI), whereby students work together, using an interdependent group contingency, to achieve a common goal (Slavin, 1980). Interdependent group contingencies, where a group reward is contingent upon the performance of the whole group, have been

used successfully to improve classroom behaviour in multiple studies, previously (Brogan et al., 2017; Gresham & Gresham, 1982; Groves & Austin, 2017; Joslyn et al., 2019; Slavin, 1991; Wright & McCurdy, 2012). CL can improve achievement scores, intergroup relations, peer acceptance, positive attitudes, and self-esteem (Kyndt et al., 2013). Veldman et al. (2020) found 6- and 7-year-old students who had participated in a CL *Success for All* program for one year, showed more positive group work behaviour and less negative group work behaviour than a control group, during a group task. In a sample of 90 female pre-intermediate college students who spoke English as a second language, Jalilifar (2010) found students in the *Student Team Achievement Divisions* CL group performed better in an English Language Proficiency post-test than students in the *Group Investigation* CL group and control group. The authors argued that the reason for the difference in results between the two CL groups was likely because the *Student Team Achievement Divisions* technique included a reward component, which likely reinforced student behaviour, whereas the *Group Investigation* technique was missing this feature. Using a variety of scales and questionnaires to measure bullying and victimisation, perceived stress, emotional problems, relatedness, and engagement, Van Ryzin and Roseth (2018) investigated the effects of CL on peer relations, victimization, bullying, and related outcomes of 1,460 seventh-grade students, using a cluster randomised controlled trial. After an average of 5.5 months between baseline and follow-up, significantly lower scores in bullying, victimisation, and perceived stress for marginalised students were found, alongside reduced emotional problems and enhanced relatedness for all students, in the CL group compared to the waitlist control group. Overall a vast array of literature has found CL to be an effective practice for student learning in all major subject areas, of multiple age-groups, across various settings (Slavin et al., 2003).

According to social interdependence theory CL generates positive interdependence, creating a cooperative environment and resulting in success when peers perceive that the only

way for group members to achieve their goals is through success of the whole group (Johnson et al., 2010). From this perspective, team building exercises prior to- and group self-evaluation after cooperative learning exercises drive success (Slavin et al., 2003). However, motivation, promoted through the use of reward systems, is critical within CL (Slavin et al., 2003). According to Slavin et al. (2003), working together towards a common goal encourages students to reinforce appropriate behaviour and learning of peers, thus promoting norms which favour academic achievement. Therefore, team building exercises and self-evaluation are insufficient in ensuring CL results in better achievement outcomes than individual tasks, without the implementation of reward systems (Slavin, 1995). This theory is supported by the results of the Jalilifar (2010) study, cited above, where students performed better in a reading comprehension test after participating in a CL group which included rewards, than the CL group without rewards and the control group. Interventions using interdependent group contingencies resulted in more significant achievement outcomes for students, according to researcher who conducted a meta-analysis on peer-mediated interventions with elementary students (Rohrbeck et al., 2003). This emphasis on reward aligns with the applied behaviour analytic view that a history of reinforcement results in an increase in behaviour (Skinner, 1968).

PMIs for behaviour management

Peers play a powerful role in the behaviour of others in the classroom (Cooper & Jacobs, 2011) and students who display disruptive behaviour can promote the same behaviour in others (Barth et al., 2004). However, alongside the teaching of other skills, there is considerable research highlighting the success of peer-mediation on classroom behaviour management (Coogan et al., 2007; Cooper & Jacobs, 2011; Grauvogel-Macaleese & Wallace, 2010; Kaya et al., 2015; Sinclair et al., 2019). Teachers cannot observe all children all of the time and will inevitably miss opportunities to reinforce appropriate behaviour (Lum

et al., 2017; Skinner et al., 2002). Instead of relying on teachers to observe and respond to student behaviour, peer-mediated interventions use the abundant resource of students, leaving teachers more time to focus on instruction (Lambert et al., 2015; Lum et al., 2019; Lum et al., 2017). Results from meta-analyses have shown that peer mediators can effectively improve behaviour of students of all ages with and without emotional and behavioural disorders (Dart et al., 2014; Kaya et al., 2015). In a conceptual replication of a previous study, Sinclair et al. (2019) investigated the effects of a reciprocal peer-tutoring programme, on disruptive and academically engaged behaviour, of a 13-year-old male student in an intervention classroom receiving Tier 2 reading supports from a general education teacher. In an ABAB design, following the initial baseline phase, during which students participated in an individual, online reading programme, the teachers, followed by students in the class were trained on implementation and use of Peer-Assisted Learning Strategies 2-6 (PALS). During intervention phases, pairs of students participated in four activities consisting of *Partner Reading, Retell, Paragraph Shrinking,* and *Prediction Relay*. Disruptive behaviour decreased and academically engaged behaviour increased during intervention phases, compared to baseline. Sinclair et al. (2019) argued that these results were likely due to changes in classroom context from self-directed learning to a more instruction-based structure, and opportunities to engage with peers and respond to instructional stimuli. The authors also discussed the function of the participants' behaviour as an explanation for the results, theorizing that disruptive behaviour often resulted in peer attention during baseline, whereas the PALS intervention provided peer attention contingent upon on-task behaviour. Similarly, PALS provided escape from nonengaging instructional demands, if escape was the function of behaviour during baseline. However, without a functional assessment, which would have determined the actual cause of behaviour, it was only possible to speculate the function of

behaviour, and the intervention may have been less successful with other functions, such as teacher attention or access to tangibles.

Social skills

Many PMIs, such as PALS and peer tutoring, for example, originally designed to improve academic achievement, can also positively influence classroom behaviour (Cooper & Jacobs, 2011; Sinclair et al., 2019). Similarly, improving student social skills has been found to enhance appropriate classroom behaviour (Gresham et al., 2004), as social skills deficits are often found in students displaying high levels of disruptive behaviour (Walker et al., 2003). In their review of six meta-analyses on social skills training, Gresham et al. (2004) found social skills training to be an effective intervention for children with emotional and behavioural disorders, improving behaviours such as aggression, internalising and externalising behaviours, and antisocial behaviour. Unfortunately, skills learned in teacher-mediated social skills training, which occurs outside of the general classroom, often do not generalise to other settings, because of a lack of exposure to natural stimuli (Stokes & Baer, 1977). However, some PMIs are designed to promote prosocial behaviour by suppressing reports of antisocial behaviour, otherwise known as “tattling”, within the classroom (Cashwell et al., 2001; Moroz & Jones, 2002; Skinner et al., 2000; Skinner et al., 2002). Skinner et al. (2002) argued that punitive behaviour management practices which result in students learning to avoid being “caught” by the teacher can create environments whereby teachers rely on student reports of antisocial behaviour of their peers (tattling). Consequently, tattling may cause students to focus on antisocial behaviour, thus ignoring prosocial behaviour, and resulting in social isolation of peers (Skinner et al., 2002). In contrast, positive peer reporting interventions, which provide students with opportunities to report prosocial behaviour of their peers can improve social interaction, peer acceptance, and social

involvement (Moroz & Jones, 2002; Skinner et al., 2002). One class-wide positive peer reporting procedure which has had positive results in the classroom is tootling.

Tootling

Rooted in ABA principles, tootling is an example of a PMI whereby students monitor and privately report instances of prosocial peer behaviour. After training on how to notice and report prosocial behaviour, students are given blank slips of paper for reporting throughout the day, which are then collected in a tootle container. The total number of tootles collected each day is publicly displayed on a chart. Teachers randomly select and read several tootles at the end of each day and provide praise and/or corrective feedback to the class. This differential reinforcement results in an increase in positive reports of behaviour and a decrease in negative reports (Murphy & Zlomke, 2014). Class-wide rewards are agreed upon by students and teachers, and issued for reaching a tootle target (Dillon et al., 2019; Lum et al., 2017; McHugh et al., 2016; Skinner et al., 2000). Overall, tootling contains three key components.

Key Components of Tootling

Positive Peer Reporting. One key component of tootling, as mentioned, is positive peer reporting. Skinner et al. (2002) argued that if children can learn to monitor and report peers' inappropriate behaviour without being taught, then they can be taught to monitor and report prosocial behaviour. In the first reported study on the reporting of peer-prosocial behaviour, kindergarten students were given the opportunity to share cooperative or friendly behaviours they had observed by their peers. Peers who had their behaviour reported received a smiley-faced badge. Cooperative acts increased and aggression decreased during the first intervention phase, compared to baseline and withdrawal, and maintained once badges were removed (Grieger et al., 1976). Positive peer reporting has continued to demonstrate success in developing prosocial behaviour (Moroz & Jones, 2002; Morrison & Jones, 2007).

However, there is a lack of reporting in the literature on the effect of positive peer reporting on general appropriate classroom behaviour (Sherman, 2012).

Interdependent group contingency. Interdependent group contingencies are another key component of tootling. Similar to their use in CL, working towards a shared goal may focus students' attention on target behaviour and promote a sense of solidarity between peers (Gresham & Gresham, 1982; Slavin, 1991). Another example of a well-researched, effective behaviour intervention which uses an interdependent group contingency to promote prosocial and appropriate classroom behaviour is the Good Behaviour Game (GBG) (Bowman- Perrott et al., 2016). The GBG was originally introduced by Barrish et al. (1969), with the main features including, students being placed into teams, teams accumulating points for inappropriate behaviour, and the team with the fewest points winning a reward. Although this initial study focussed on punishing inappropriate behaviour (by awarding points, when the goal is to have the fewest points), others have successfully modified the game to focus on appropriate behaviour (Wahl et al., 2016; Wright & McCurdy, 2012). One limitation of the GBG is that in most studies there is a reliance on the teacher noticing behaviour (Mitchell et al., 2015; Rubow et al., 2018; Wahl et al., 2016; Wright & McCurdy, 2012), although it has demonstrated adaptability in multiple ways (Bowman-Perrott et al., 2016).

Public posting. Public posting is a method of providing numerical or pictorial feedback, for example, via a medium within the environment, such as a poster board or whiteboard (O' Handley et al., 2020). Skinner et al. (2000) initially included public posting of progress as a stimulus to prompt students to tootle. Although most studies assessingthe effects of public posting have included it within packaged interventions, such as tootling, O' Handley et al. (2020) found public posting had an additive effect on academically engaged behaviour, but not disruptive behaviour, in secondary-aged students. Overall, alongside other PMIs, tootling, with interdependent group contingency, and public display of progress may

ease teacher-burden, potentially increasing the feasibility and social validity of the intervention (Lum et al., 2019; McHugh et al., 2016).

The tootling literature

Initially, an interdependent group contingency, alongside public display of progress was used to increase students' reports of peer prosocial behaviour (Skinner et al., 2000), and tootling has since been found to successfully improve multiple areas of classroom behaviour (Cihak et al., 2009; Dillon et al., 2019; Kirkpatrick et al., 2019; Lambert et al., 2015; Lipscomb et al., 2018; Lum et al., 2019; Lum et al., 2017; McHugh et al., 2016). Originally, using an ABAB design, Skinner et al. (2000) measured number of tootles as a function of an interdependent group contingency and public display of progress, among a general education class of twenty-eight 9- to 10-year-olds. The authors sought to design a procedure which increased awareness and reinforcement of classroom prosocial behaviour.

Students were initially trained how to tootle on their peers and, during baseline, note cards were taped to their desks and students were encouraged to place completed tootle cards, describing peers' prosocial behaviour, into a tootle box on the teacher's desk. As the dependent variable, total number of tootles was counted by the researchers at the end of each day. An interdependent group contingency and public display of progress was introduced during the first intervention phase, whereby a tootle target of 100 was set and students were informed of a class reward of 30-minutes extra recess for reaching that goal. After the tootles were counted each day, the cumulative total was added to the public display. This visual stimulus allowed students to see their daily progress. Once this target was met, the reward was provided that day. The following day, the target was increased to 150 tootles and students received a different reward for reaching the target. During a return to baseline, the public display was removed, along with the interdependent group contingency and students

were informed they could still place tootles in the tootle box but there would be no reward for doing so. The final intervention phase included a third reward and a target of 150 tootles.

Number of tootles collected was highly variable during the initial baseline and treatment phases, followed by an immediate decrease to near zero, then increase in tootling, during the second withdrawal and treatment phases, respectively. The authors speculated that this initial variability may have occurred because of an unrelated group-contingency punishment procedure, implemented by the principal, alongside the first two phases. Regardless, behaviour was more stable during the third and fourth phases. Although no social validity scores were collected, the authors noted that the teacher continued to implement the programme one week after the research ended, which could suggest the teacher believed it to be a successful intervention. However, no follow-up data were collected to examine the veracity of this claim. Alongside a paucity of social validity scores and follow-up data, there were no reports of treatment integrity. This may have been because the researchers completed the majority of the components. Nevertheless, failing to report treatment integrity creates difficulty in determining the accuracy and consistency with which the intervention was implemented (Gresham, 2005). Still, it was concluded that the interdependent group contingency, alongside public posting was successfully used to increase reports of prosocial behaviour in the classroom, although replications of results would be beneficial to the research.

In a replication and extension of the Skinner et al. (2000) study, Cashwell et al. (2001) used similar methods to measure the number of tootles reported by 7- to 8-year-old students in response to direct instruction, an interdependent group contingency, and public display of progress. During intervention phases, students earned 20-minutes extra recess for reaching 100 tootles, a field trip to a special playground for reaching 150 tootles, then movie and popcorn for reaching 200 tootles, with the counter re-setting each time a target was met.

Results were variable across phases one, two, and three, although clear changes in the rate of tootling after all phase changes were evident, suggesting experimental control and positive effects of the group contingency with public progress display. The authors argued that the intervention worked to increase reporting of prosocial behaviours and the study appeared to have good social validity, as the teacher continued to implement the procedures during the final two-and-a-half weeks of the school term. However, social validity, follow-up data, and treatment integrity scores were again absent, threatening the overall validity of the study. Additionally, only tootles which could be read and understood by the researchers were counted towards the target, meaning tootles written by students with poor handwriting may rarely have been counted, if at all, thus potentially confounding results.

Limited but promising results were found in these studies, in relation to tootling, but neither measured the effects of tootling on other classroom behaviour. Cihak et al. (2009) extended the previous research by investigating the effects of tootling on disruptive classroom behaviour of 19 third-grade students with and without disabilities, in an inclusive classroom, and assessing teacher acceptability of the procedure as a classroom-based positive behaviour support.

In an ABAB research design, the teacher recorded disruptive behaviour on a bracelet made from construction paper, by adding a mark next to the initial of the appropriate student, using event recording. During baseline, the teacher conducted the class typically. The students were trained on tootling by the teacher before the first intervention phase. Then index cards, for recording tootles, were placed on student desks and collected throughout the day. A tootle target of 75 was agreed upon by the researchers and teacher, and reaching it earned the students a predetermined group reward, such as 20-minutes extra recess. Total daily tootles were cumulated until the target was reached. Treatment integrity scores were calculated to assess reliability of the teacher’s observations and ranged from 86% to 100%.

A decrease in disruptive behaviour during the two intervention phases was found, compared to baseline and withdrawal phases, with 100% non-overlapping data points. Additionally, there were no occurrences of disruptive behaviour during the last three days of the intervention. The teacher rated the intervention socially acceptable on the Intervention Rating Profile-15, which is a 15-item Likert scale questionnaire which measures how acceptable the participant deems an intervention to be. Overall, the tootling intervention, using interdependent group contingencies and public display of progress, was successful in reducing classroom disruptive behaviour of students with and without disabilities. In the introduction to their research, the authors discussed the possibility of teachers being unable to attend to all classroom behaviour due to teaching demands. Therefore, it is likely that using the teacher to record student behaviour may not have produced accurate results and using external observers may have been more appropriate.

Following on from this research, Lambert et al. (2015) aimed to replicate the above research, using trained observers to measure changes in levels of disruptive and appropriate behaviour in two fourth and fifth grade classrooms, because of a tootling intervention. Disruptive and appropriate behaviours were recorded using momentary time sampling in an ABAB design with a multiple baseline across classrooms. Like previous studies, during intervention phases, class tootle targets were set, and class rewards issued on reaching that target. Targets increased following goal-achievement. Unlike previously, the authors of this study recorded data during a two-week follow-up, after teachers were given the option of continuing or not.

While the previous researchers relied on visual analysis alone to interpret results, Lambert et al. (2015) calculated effect sizes of disruptive and appropriate behaviour, using Nonoverlap of All Pairs (NAP). They found a decrease in disruptive behaviour and an increase in appropriate behaviour across both classrooms during the intervention and during

follow-up, compared to baseline, with moderate to strong effect sizes in classroom A and strong effect sizes in classroom B. Teachers rated the intervention acceptable and both continued its use during follow-up. No generalisation across settings data was recorded, making it difficult to understand any effects of tootling outside the classroom in which it was taught.

Following on, McHugh et al. (2016) extended the research by assessing the effects of tootling on disruptive and academically engaged behaviour on class-wide, as well as individual target student behaviour, in lower elementary general education classrooms, using a tootle target which could be achieved daily. Participants were 64 second- and third-grade students from three classrooms across two schools. One target student, with high levels of disruptive behaviour, was identified from each classroom, none of whom received special education services. Many procedures were similar to that of the Lambert et al. (2015) study, with one major variation being the criterion number of tootles required to reach the class target, which was set at 25 or 30 for each class, based on the possibility of each child writing one or two tootles each, and the target being achievable daily.

Results were consistent with previous studies, showing lower-level disruptive behaviour and higher-level academically engaged behaviour class-wide and for individual target students during the tootling phase, compared to baseline phases. Effect sizes were moderate to large across phases class-wide and for target students. Changing the criterion number of tootles needed to reach the target did not attain different results from previous studies using higher tootle targets, although a direct comparison of the two criterion-types was not performed, making conclusions about the differences in effect difficult.

A downward trend in disruptive behaviour was found in one classroom and an upward trend in academically engaged behaviour was found in all three classrooms during the withdrawal phase, before moving into the final intervention phase. These data could be an

example of response maintenance, whereby behaviour continues once intervention materials have been removed (Stokes & Baer, 1977). However, this was not discussed by the authors. Furthermore, it is possible that tootling may have affected teacher behaviour which, in turn, may have affected student behaviour, even once the intervention was removed. For example, the teachers may have become more aware of appropriate student behaviour during tootling, thus using more praise statements. This behaviour change may have maintained during withdrawal, thus affecting student behaviour. Regardless, behaviour further improved during the second intervention phase and overall, tootling was an effective procedure for improving classroom behaviour.

Lum et al. (2017) extended the research further by assessing the effects of tootling on disruptive and academically engaged behaviour in three classrooms of high school students, although the age of the students is unknown. Procedures mostly resembled those of previous studies, but one age-appropriate amendment was made by asking students to vote on an appropriate name for the procedure. As with the McHugh et al. (2016) study, disruptive behaviour was the primary dependent variable, meaning phase change decisions were based on this variable and not academically engaged behaviour.

Disruptive behaviour decreased and academically engaged behaviour increased during intervention phases in all three classrooms, compared to baseline, but results were not maintained during a 2-week follow-up. The teachers rated the intervention moderate to high on the Behaviour Intervention Rating Scale (BIRS), each strongly agreeing that they would use tootling again. However, none were using the intervention during follow-up, contradicting these results. Conversely, one teacher anecdotally reported feeling less stressed at the end of each day during the intervention. Collecting data on reasons why the teachers were not using the intervention may have provided some insight into these findings.

Measuring and comparing data on teacher attitudes and behaviour during all phases of the study could also contribute to the outcomes of the research.

Addressing gaps in the literature, Lipscomb et al. (2018) compared the effects of ClassDojo (an online behaviour management tool designed to allow teachers to monitor and provide real-time feedback for student behaviour), and tootling plus ClassDojo on classroom behaviour of seven adult university students with intellectual disabilities. Participants were 19 - 24-year-old adults in their first year in a full immersion Comprehensive Transition Classroom. The students were referred because of concerns about levels of disruptive behaviour. Therefore, disruptive behaviour was the only dependent variable.

After collection of initial baseline data, within an alternating treatment design, one of two interventions, or a control, was randomly assigned to the class each day. During baseline and control phases, the instructors conducted the class as normal. During ClassDojo phases, instructors awarded points to students for appropriate classroom behaviour and rewards were provided when the target number of reports was reached. During tootling plus ClassDojo phases the students were allocated times throughout the lesson to record their own observations of appropriate peer behaviour, using ClassDojo. The same criteria were used for reward-delivery as ClassDojo phases. Data was collected using a frequency count of each participant's behaviour, before results were combined to give a total for the class.

Using visual analysis and a nonoverlapping of all datapoints procedure to analyse the data, researchers found that both intervention conditions reduced disruptive behaviour, with strong effect sizes, compared to control and baseline conditions. This adds to the literature regarding the use of ClassDojo in the classroom, confirming that a peer-mediated element can be added, to ease teacher-burden. However, without comparing tootling without ClassDojo, it is impossible to establish any differences between tootling with- and without a technological component, such as ClassDojo.

Dillon et al. (2019) investigated the effects of a ClassDojo-enhanced version of tootling on disruptive behaviour and academically engaged behaviour of 35 students in three fifth-grade classrooms. In an ABAB design, following baseline, teachers trained students how to identify and report peer appropriate behaviour (tootle) using ClassDojo. Following a 20-minute instruction period, students were invited to spend 5 seconds in a tootle booth uploading tootles. Consistent with previous tootling studies, students worked together to earn a class reward. Public posting occurred in the form of an update on an interactive whiteboard, automatically updated as part of the ClassDojo software. These updates therefore provided immediate feedback, potentially increasing effectiveness of the intervention, compared to traditional tootling, according to reinforcement theory, assuming feedback itself is reinforcing. However, there was no comparison between the two tootling types to test this theory.

Levels of disruptive behaviour decreased and academically engaged behaviour increased during intervention phases, compared to baseline and withdrawal phases, with strong effect sizes for disruptive behaviour and moderate to strong effect sizes for academically engaged behaviour. Although social acceptability scores were moderate to strong, effectiveness scores were consistently lower than other scores. The authors argued this may have been a result of the teacher's perceptions regarding the sustainability of the intervention, especially because increases in disruptive behaviour and decreases in academically engaged behaviour were found during withdrawal. Behaviourally, this unwanted behaviour could be explained as an extinction burst, whereby a temporary increase, followed by a decrease in an undesirable response occurs after removal of a reinforcer. Programming for response maintenance and conducting maintenance probes to measure whether behaviour changes were sustained after removal of the intervention may have helped to increase effectiveness scores.

Similarly to all previous studies which measured the effect of tootling on classroom behaviour, Dillon et al. (2019) failed to record data on number of tootles produced during intervention phases. The authors argued that earlier results, from Cashwell et al. (2001) and Skinner et al. (2000) indicated that tootling increased during intervention phases. However, recording the daily tootle count would have provided valuable data on treatment integrity.

In an ABAB design, Kirkpatrick et al. (2019) investigated the effect of a modified tootling intervention on antisocial and/or disrespectful behaviour of four 8-to 10-year-old students in a voluntary afterschool programme for at-risk students. Instead of a class reward contingent upon reaching a known tootle-target, the target was a random, unknown number of class-peers who received tootles about them. This alteration was designed to encourage all students to participate in prosocial behaviour. Other procedures were similar to previous tootling studies, except that the researchers trained students on tootling, because the teachers were volunteers. No social acceptability scores were measured.

Using visual analysis and effect size measures it was found that the antisocial and/or disrespectful behaviour, towards peers and staff, decreased as a result of the tootling intervention. However, there was no measure of prosocial behaviour. Similarly, although student behaviour towards teachers was recorded, teacher behaviour was not. It is possible that teacher behaviour, as a result of the intervention, may have prompted student behaviour, affecting results. Nevertheless, the results added to previous research indicating tootling can decrease antisocial and/or disrespectful behaviour in a classroom setting.

Lum et al. (2019) also modified the tootling intervention to assess the effects of tootling with a randomised independent group contingency on disruptive behaviour and academically engaged behaviour in three general education high school classrooms. Seventy-two tenth-, eleventh-, and twelfth-grade, students participated. Procedures were typical of previous studies, except a randomized independent group contingency was used, whereby

rewards were delivered to three randomly chosen students who had tootles written about them and two who wrote tootles, each day.

Decreases in disruptive behaviour and increases in academically engaged behaviour were found during intervention phases, compared to baseline and withdrawal. Teachers and students both rated the intervention acceptable using the modified BIRS and modified Children's Intervention Rating Scale (CIRP). However, no generalisation data were collected so it is unknown if any behaviour changes occurred in other settings as a result of the intervention or were maintained over time. Again, no data were recorded on number of tootles submitted or teacher behaviour, creating difficulty in assessing any functional relation between these factors and student behaviour. Although positive results were reported in this study, a previous comparison of interdependent and dependent group contingencies reported better outcomes using the interdependent group contingency (Gresham & Gresham, 1982).

In an effort to investigate the effects of a single component of the tootling procedure, the written component, Derieux (2019) compared the effects of traditional tootling to a writing procedure, and no-treatment control, on academically engaged behaviour and disruptive behaviour of students in three, nine- to twelfth-grade classrooms. An alternating treatment design was embedded within a multiple baseline design, across three classrooms. The traditional tootling condition consisted of the typical components found in other studies. The writing condition was similar to the traditional tootling condition, but instead of writing prosocial behaviour of peers, the students wrote about two things they had learned that day. In the control condition the students again wrote about what they had learned, however, in this condition, all reinforcement was removed.

The author reported mostly small to moderate effect sizes for academically engaged behaviour, disruptive behaviour, and passive off-task across all three conditions, in all classrooms and concluded that the writing component was not the controlling variable

causing behaviour change. However, although the tootling and written phases both included rewards for reaching predetermined goals, goals were only reached six times throughout the study and one of those occasions was during the no-treatment control, which included no reward. It is possible that the goal of two notecards per student was set too high, potentially confounding results.

Goss (2019) examined the effects of tootling without goals and external rewards on the negative behaviour of 24 lower-elementary students, aged 6- to 9-years-old, in a Montessori classroom, using an AB design. Students completed self-assessment forms reporting the behaviour of themselves and their peers during group lessons, before the researcher defined and discussed correct classroom etiquette with them. Examples of correct etiquette included raising hands before talking, only talking when it was the student's turn to talk, and looking at the individual who was talking. Students then role-played appropriate and inappropriate classroom etiquette. Following, the researcher read a book to the students about prosocial behaviour, then another about telling others when they do nice things, before training on tootling began. After placing a tootle in the container, students were asked to place a flower on a picture of a tree, which represented peace blooming in the classroom, in-line with one of the books they read together. Tallied scores of negative behaviours were used for data analysis, alongside comparisons of the self-assessment forms, which were completed prior-to and after the intervention phase.

Goss (2019) reported no relationship between tootling and negative behaviour, with increases in some behaviours and decreases in others, though analysis of the self-assessment showed that students believed their behaviour had improved. Although no data was collected on teacher behaviour, after zero levels of tootling during the first few days, the researcher and teacher agreed that the teacher would prompt students to tootle at the beginning and end of each lesson.

Caution is needed when interpreting these results and comparing to the previous tootling literature. Alongside the intervention, the teacher also introduced a separate social skills training package, which could have impacted results. Negative behaviour was not operationally defined, making comparisons to previous studies difficult. Negative behaviours were also selected using behaviours observed in different students from the previous year. For example, tattling, which was considered a problem in the previous cohort, was used as a measure, even though no tattling occurred during baseline. Regardless of these results, the author concluded that tootling without goals and rewards had no effect on negative classroom behaviour.

Collectively, results show tootling, alongside an interdependent group contingency and public display of progress can increase reports of peer prosocial behaviour in the classroom (Cashwell et al., 2001; Skinner et al., 2000) and decrease levels of antisocial and/or disrespectful behaviour (Kirkpatrick et al., 2019). Additionally, levels of disruptive behaviour have been found to decrease and levels of academically engaged behaviour increase, as a result of tootling alongside both interdependent and independent group contingencies and public display of progress, for students of multiple ages, with- and without intellectual difficulties, in general- and special education classrooms (Cihak et al., 2009; Dillon et al., 2019; Kirkpatrick et al., 2019; Lambert et al., 2015; Lipscomb et al., 2018; Lum et al., 2019; Lum et al., 2017; McHugh et al., 2016). Furthermore, moderate-to-high social validity scores among teachers and students indicate that teachers and students alike perceive tootling to be an acceptable, effective, and efficient invention for classroom behaviour (Lipscomb et al., 2018; Lum et al., 2019; McHugh et al., 2016). Nevertheless, there are still gaps in the literature which need to be explored. For example, the effects of tootling on teacher behaviour, as well as any maintenance effects are yet to be examined.

Teacher behaviour

Teacher behaviour plays a critical role within the classroom, contributing towards aspects such as teacher-student relationships and overall student motivation (Bonik, 2018), peer relationships (Hendrickx et al., 2017), student wellbeing (Van Petegem et al., 2008), on-task behaviour (Chalk & Bizo, 2004), student engagement and academic achievement, (Harbour et al., 2015), and reading motivation (Naeghel et al., 2014). Much of the literature on teacher behaviour measures its effects on student behaviour (Bonik, 2018; Hendrickx et al., 2017; Royer et al., 2019; Skinner & Belmont, 1993; Spilt et al., 2016; Weyns et al., 2017) and academic outcomes (Kyriakides et al., 2009; Panayiotou et al., 2014; Ritzema et al., 2016; Zhu et al., 2018), with some studies describing how to improve teacher behaviour for the benefit of the classroom (Allday et al., 2012; Antoniou & Kyriakides, 2013; Schaffer et al., 2017; Sedova, 2017). Similarly, many classroom interventions are designed to either improve student behaviour or improve teacher behaviour and student academic outcomes. For example, using a modified multiple baseline design across four teachers and seven students with, or at risk of, emotional/behavioural disabilities, Allday et al. (2012) investigated the effects of a teacher training package, designed to increase behaviour specific praise, on teacher and student behaviour. Increases in behaviour specific praise and student on-task behaviour, and decreases in the number of corrective statements were found during intervention phases, compared to baseline. However, though successful, training packages, such as this often require a lot of time and resources, alongside being intrusive in the classroom (Eaves et al., 2020). Whilst interventions such as this are designed to target teacher behaviour, few studies report teacher and student behaviour as a result of behavioural interventions, designed to target student behaviour.

Historically, teachers have tended to use more praise statements in response to academic achievement and reprimands in response to behaviour (Harrop & Swinson, 2000;

White, 1975). With the focus shifting to a positive approach to classroom behaviour, it is important to find successful ways of increasing positive interactions between teachers and students. Using praise statements in response to appropriate behaviour reinforces that behaviour and can increase appropriate behaviour and decrease disruptive behaviour (Sutherland et al., 2000)

Although professional development and training packages designed to increase behaviour-specific praise, such as the one described by Allday et al. (2012) can provide successful classroom management strategies, they are often not sufficient in maintaining positive effects (Sawka et al., 2002). Sutherland et al. (2000) investigated the effects of an observation-feedback intervention on rates of teacher behaviour-specific praise provided to students with emotional and behavioural disorders. They found that behaviour-specific praise increased during the observation feedback phases but was not maintained during the withdrawal phase, indicating a functional relationship between praise and feedback. These results suggest that continuous feedback may be needed to maintain high levels of behaviour-specific praise. Rubow et al. (2018) argued that manipulating classroom contingencies to encourage appropriate behaviour could improve teacher behaviour without the need for interventions involving direct feedback. In other words, an intervention which improves student behaviour may, subsequently, increase levels of teacher praise, without need for continuous feedback on teacher behaviour.

Many studies designed to improve classroom behaviour introduce an intervention designed to improve student behaviour, then measure the effect of that intervention on the target behaviour (Camacho et al., 2014; Riley et al., 2011; Sinclair et al., 2019; Trussell et al., 2018). Although it has been well established that teacher behaviour plays an essential role within the classroom (Chalk & Bizo, 2004; Harbour et al., 2015; Naeghel et al., 2014), limited studies report the effects of behavioural interventions on student and teacher

behaviour, simultaneously. An increase in teacher praise statements and decreased reprimands during interdependent group contingencies have been found (Wills et al., 2014). It is possible that, as students reinforce appropriate behaviour of their peers to achieve a common goal, and appropriate behaviour increases (Slavin et al., 2003), more opportunities are generated for teachers to praise, with less demand for reprimands. Lannie and McCurdy (2007) investigated the effects of the GBG on the behaviour of one teacher and 22 students in a first-grade classroom, using a withdrawal design. They found student disruptive behaviour decreased, while on-task behaviour increased during intervention phases, compared to baseline and withdrawal. However, teacher praise statements remained at almost zero levels across phases, while negative statements tended to decrease when disruptive behaviour decreased. The authors argued that disruptive behaviour is more likely to occasion a response from the teacher than passive, on-task behaviour. However, there was only one teacher participant in this study and phase changes were determined by stability in disruptive student behaviour only. Elswick and Casey (2011) implemented the GBG in a first-grade general education, inclusive classroom, with one teacher and 20 students. Teacher and student behaviour was measured using an AB design and increases in on-task- and decreases in disruptive student behaviour were found, along with increases in behaviour-specific praise statements and decreases in reprimands, during intervention phases, compared to baseline and withdrawal. Although these results indicate that the intervention had a positive effect on both student and teacher behaviour, the AB design limited interpretation of results. Rubow et al. (2018) extended previous research by investigating the effects of the GBG on student and teacher behaviour across two alternative learning centre classrooms. Participants were 22 students ranging from fourth- to eighth-grade, two teachers, and two paraprofessionals. A reversal design was used in one classroom and a multiple baseline across settings design in the other. During implementation of the GBG, disruptive behaviour decreased, while the

teachers' use of praise statements increased compared to reprimands. The varied results from these studies show that more research is needed to investigate the effects of successful behaviour interventions on teacher and student behaviour.

Although some research has examined these effects, there is a lack of literature investigating effects of peer-mediated interventions, designed to improve student behaviour, on teacher behaviour. In their meta-analysis investigating single-case research on peer-mediated interventions, Dart et al. (2014) found a lack of reporting on teacher behaviour in the form of praise, feedback, or approval. It is possible that peer-mediated interventions, such as tootling, could create an environment whereby teachers are naturally alerted to positive prosocial behaviour of students, thus resulting in an increase in praise statements and a decrease in reprimands.

While the evidence suggests that tootling can be an effective and efficient intervention for managing classroom behaviour, there is no existing literature on the effect of tootling on teacher behaviour. Kirkpatrick et al. (2019) suggested that tootling behaviour may have resulted in an increase in publicly labelled praise, from teachers, due to a newly increased awareness of prosocial behaviour and/or a desire to prompt students to report on these behaviours. Cashwell et al. (2001) reported anecdotal evidence of a teacher's surprise at the amount of prosocial behaviour the students participated in during baseline. The authors speculated that an increase in awareness, such as this, may have elicited an increase in publicly labelled praise, thus resulting in more tootling and better outcomes within the classroom. Henington and Skinner (1998) suggested that teachers may focus so much on inappropriate behaviour that they are unaware of how frequently appropriate behaviour occurs. Additionally, Skinner et al. (2000) suggested future researchers examine the effects of tootling on teacher behaviour, as increased praise levels may result in improved classroom relationships.

Generalisation/maintenance

A defining characteristic of ABA, and therefore of interventions designed to meet the PB4LSW criteria, is generalised behaviour change (Baer et al., 1968). According to Baer et al. (1968) generalisation occurs when a behaviour which has been trained in one context also changes in another context, such as a different time (also known as maintenance) or environment, or if an intervention results in behaviour change, other than the target behaviour. For example, Moore et al. (2001) investigated the effects of a self-management intervention incorporating elements of self-recording and goal setting, to help train generalisation in a natural setting, on three 8-year-old boys with limited on-task behaviour. The authors reported maintenance of improved behavioural outcomes due to the intervention during class-time once the intervention was removed. The teachers also reported generalisation effects of two students who showed behavioural improvements during other class-times, where the intervention was not implemented. This study highlights the way generalisation can be trained, as in the maintenance of behaviour reported once the generalisation intervention was removed, and can also occur as a natural consequence of the intervention (Baer et al., 1968). Without generalisation, any behaviour change would be limited to the context in which it was learned.

Taught behaviours which contact natural environmental contingencies are more likely to be maintained than behaviours that do not encounter these contingencies (Stokes & Baer, 1977). For example, adult-mediated social skills interventions have successfully taught social skills to students, but these skills often do not generalise into the classroom (Cashwell et al., 2001). This is likely because the reinforcement (probably praise) provided during training sessions does not continue in the classroom, thus highlighting the importance of teachers "catching students being good". Therefore, interventions introduced in the classroom are more likely to increase response maintenance than those taught elsewhere. Results from

multiple peer-mediated interventions have shown success in generalisation and maintenance of skills (Bambara et al., 2016; Grauvogel-Macaleese & Wallace, 2010; Schmidt & Stichter, 2012; Thiemann & Goldstein, 2001). In a multiple baseline across participants design, Bambara et al. (2016) investigated the effects of a peer-mediated intervention on the conversational skills of high school students with Autism Spectrum Disorder, by teaching typically developing peers strategies for engaging in conversation with focal students. Text-cues and direct instruction for focal students were also used to increase flexibility and aid with the natural course of conversation. Although no follow-up trials were performed, once training ended, focal students showed continued improvements of conversational abilities with peers. An important aspect of this research was that during post-test observations, communicative peers remained the same as during the training phases and there was continued use of text-cues by focal students. This meant that potential discriminative stimuli for appropriate responding remained during the post-test environment, likely contributing to response maintenance. In another study, after completion of a functional analysis showing off-task behaviour was a function of peer attention, Grauvogel-Macaleese and Wallace (2010) investigated the effects of a peer-mediated intervention on off-task behaviour of three students with attention deficit hyperactivity disorder. Peers were trained to provide attention to participants contingent on on-task behaviour. Decreases in off-task behaviour during peer-attention phases, compared to baseline, were found. Off-task behaviour remained at near-zero for one student who participated in follow-up trials one month later. Similarly to the Bambara et al. (2016) study, attention was provided by the same peer during follow-up. By training in the natural environment and using peers and materials during training trials that remain once interventions cease, participants continue to encounter reinforcement in the natural environment and improved rates of behaviour remain (Baer, 1999).

Although some interventions do result in response maintenance, not all will do so without specific programming (Stokes & Baer, 1977). While tootling is a peer-mediated intervention performed in the classroom, unlike the studies described above, appropriate behaviour may not be under control of natural environmental contingencies, such as peer responding. This could explain a lack of response maintenance reported in the literature. It is possible that, instead, the tootle apparatus and, more likely, the contingencies in place that result in reinforcement, are the discriminative stimuli controlling behaviour. Therefore, once these discriminative stimuli are removed, there are no environmental stimuli indicating reinforcement is available, resulting in a decrease in appropriate behaviour. In their research on the maintenance effects of the GBG, Lynch and Keenan (2018) anecdotally described a situation whereby a student raised her hand to get the teacher's attention, then vocalised that she remembered the class was not playing the game that day and proceeded to lower her hand and call-out to the teacher instead. Consequently, the authors argued that behaviour was under control of the discriminative stimuli related to the contingencies developed by the game and that maintenance needed to be programmed to allow differential reinforcement of appropriate behaviour in the natural environment. This aligns with Stokes and Baer (1977) that sometimes generalisation must be programmed. It is also feasible that student behaviour could be under control of the discriminative stimuli related to teacher behaviour, such as praise and corrective statements. Therefore, if praise statements increase during an intervention and corrective statements decrease, it is possible that there will be less need for maintenance programming, especially if praise functions as a reinforcer for appropriate behaviour, which it likely does, according to previous research which has found appropriate behaviour to increase because of increased praise (Chalk & Bizo, 2004).

Sequential Withdrawal

Sequential withdrawal describes the withdrawal of each component of a multi-component treatment, sequentially, during consecutive phases until all components have been removed (Martin & Rusch, 1987). Dillon et al. (2019) and Lum et al. (2017) referred to sequential withdrawal when they suggested gradual fading of the tootling intervention as a way of allowing more natural consequences to exert control over behaviour.

Previous studies have found success in fading behavioural interventions. In their study investigating the effects of self-management training to improve social skills of a 9-year-old girl with Autism Spectrum Disorder Liu et al. (2015) used video modelling to help teach the participant three target behaviours. A multiple baseline across behaviours design was used with fading of self-management recording materials once a steady and high rate of target behaviour was reached. Generalisation probes across settings and follow-up observations were also completed. All three behaviours improved, and generalised across settings, as a result of the intervention. Rates of responding remained high during fading phases and continued during one-month follow-up trials. The authors concluded that the intervention was successful at improving social skills of a 9-year-old girl with ASD, with evidence of generalisation and maintenance. Fading has also successfully contributed to the maintenance of improved classroom behaviour. Petursdottir and Ragnarsdottir (2019) investigated the effects of function-based support plans, with fading of token systems, on disruptive behaviour and academically engaged behaviour of three 7- to 8-year-old male students. Using a multiple baseline across participants design, the researchers set a goal of independent functioning for all participants. Behaviour improved because of the function-based support plans and improvement rates remained for all students during fading. Once observations and data recording stopped, anecdotal evidence suggested that the teachers continued to apply some of the components of the original support plans such as clear

instructions and positive feedback. This supports the idea that fading helps increase exposure to natural contingencies. Together, these studies show that fading can be a successful approach to removing an intervention whilst maintaining learned behaviour.

Maintenance during Tootling

McHugh et al. (2016) argued that tootling can be used as a proactive PB4LSW primary tier support or as a secondary tier support in classrooms requiring extra help. The positive results showing improvements of student behaviour from the tootling literature suggests that tootling could be implemented as a continuous primary tier, proactive measure to ensure a positive classroom environment, as results indicate that tootling improves classroom behaviour (Cihak et al., 2009; Dillon et al., 2019; Lambert et al., 2015; Lipscomb et al., 2018; Lum et al., 2019; Lum et al., 2017; McHugh et al., 2016). However, the evidence regarding tootling as a secondary tier support is unclear. Secondary tier supports are designed to offer a minority of students the tools they need for a successful return to the primary tier (Lane et al., 2013). Although tootling, as a class-wide intervention, has been successful at improving individual student behaviour (Kirkpatrick et al., 2019; Lipscomb et al., 2018; McHugh et al., 2016), thus aligning with this description, there is little evidence of response maintenance. In fact, most tootling studies using an ABAB withdrawal design found that behaviour levels returned to those similar to baseline once the intervention was removed (Cihak et al., 2009; Dillon et al., 2019; Lambert et al., 2015). Only McHugh et al. (2016) reported results different to this, which were not discussed in terms of potential response maintenance and no maintenance data were reported. Furthermore, two studies which involved follow-up trials found contradicting results (Lambert et al., 2015; Lum et al., 2017). In the Lambert et al. (2015) study, results during follow-up resembled those of the intervention phases, whereas in the Lum et al. (2017) study, results during follow-up were like baseline and withdrawal phases. One major difference between these studies,

which could explain the difference in results, was that teachers in the Lambert et al. (2015) study chose to continue to use the tootling intervention, whereas teachers in the Lum et al. (2017) study did not.

While programming for any type of generalisation has not been discussed in the current tootling literature, a failure to maintain behavioural improvements has been reported. Therefore, investigating fading as an approach to programming response maintenance could add value by providing a method for removing the tootling intervention without losing the positive effects it has on classroom behaviour.

Purpose of this study

Teachers and students alike need successful behaviour management strategies in the classroom. Previously preferred punitive techniques are unsuccessful and detrimental to well-being and development. PB4LSW uses an applied behaviour analytic approach to incorporate evidence-based practices, preventing inappropriate classroom behaviour by focussing on reinforcing appropriate classroom behaviour. Although some evidence-based strategies, such as token economies, have successfully improved classroom behaviour, it has been argued that there is a lack of evidence to support them as best practice in the classroom and that they put unwanted pressure on teachers. Peer-mediated interventions reduce pressure on teachers and have been found successful at improving a variety of classroom skills. One specific peer-mediated intervention is CL, which introduces an interdependent group contingency and reward system to promote motivation to work together towards a common goal. Alongside improving student skill levels, peer-mediated interventions have improved appropriate classroom behaviour, allowing teachers more time to focus on instruction. Positive impacts on classroom behaviour have also been found as a result of social skills training packages, but changes are rarely maintained outside of the instructional setting. Tootling, alongside an interdependent group contingency and public display of progress overcomes this issue by

teaching students to recognise and report peers' prosocial behaviour, thus increasing the probability of this behaviour in future. Tootling has been used to successfully improve a variety of behaviours among a variety of individuals. However, there are gaps in the literature, such as a lack of data on the effects of tootling on teacher behaviour, which need to be addressed.

Therefore, my study will examine the effects of tootling on class-wide student and teacher behaviour. The research questions were:

1. Will a tootling intervention, implemented in year 1 and year 5/6 primary-school classrooms, increase appropriate student behaviour, class-wide?
2. Will a tootling intervention, implemented in year 1 and year 5/6 primary-school classrooms, decrease student disruptions, class-wide?
3. Will a tootling intervention, implemented in year 1 and year 5/6 primary-school classrooms, increase teacher praise statements?
4. Will a tootling intervention be rated an acceptable classroom intervention by classroom teachers and students?

Method

Participants and Setting

The participating school was a PB4LSW mainstream, government school, located in a suburb of one of New Zealand's major cities and had a decile rating of 6, indicating the socio-economic status of the area. Decile numbers are ranked from 1 (indicating the lowest socio-economic status) to 10 (the highest socio-economic status) (Ministry of Education, 2020). There were approximately 555 students enrolled at the school at the time of data collection, of which 24% identified as being Māori, 3% identified as being Pasifika, 8% identified as Asian, 60% identified as New Zealand European, and 8% identified as "other".

Within this school, three classes of students and two teachers from those classes participated in the current study. Two of the classes were "home" classes of students who were together, with their teacher, for all, or most, of the day. The third class was a mathematics class, with students with a similar skill level, from various other home classes in that year-group, who were together for 55-minutes each day. Due to these differences in classes, they are referred to as "Groups" throughout the study. A teacher-selected sample of six or seven students, representative of each class, was observed directly in a natural observation setting and measures from these student-samples were used to calculate aggregate scores across students within each group. In Group 2, during observations, it became apparent that the intervention was having a considerable impact on one student under observation. During conversations with this student's teacher, it was discovered that the student was at-risk of emotional and behavioural disorder and was receiving tertiary tier support in the form of assistance from a Resource Teacher: Learning and Behaviour. As a result of these findings, the student became an individual case study, and his on-task behaviour was recorded separately for analysis.

Teacher 1 was female, 30-years-old, with 9 years of teaching experience and 4 years teaching at the participating school. Teacher 2 was female, aged 31-years-old, with 10 years of teaching experience and 4 years teaching at the participating school.

Group 1 was a Year 5/6 class (aged 10- to 11-years-old) with 31 students. Teacher 1 taught this group and is referred to as "Teacher 1A" for the purpose of this study. Observations were conducted during writing, in the first half of the school's first teaching block.

Group 2 was a Year 1 class (aged 5- to 6-years old) with 21 students. "Teacher 2" taught this group. Observations were conducted during reading, in the first half of the school's second block, immediately following morning tea break. The first 10-minutes of this block was used for mindfulness, before reading began. The individual student was a member of this group and was a 6-year-old Māori male. For the purpose of this study, this student was given the pseudonym, "Nikau".

Group 3, the mathematics class, had 25 students from Years 5 and 6 (aged 10- to 11-years-old). Five students in this group were also in Group 1. To place students into mathematics groups, students across the entire year 5-to 6-year group were ranked by teachers, according to their skill level. The participating group was placed at the bottom of the rank, indicating a math skill level of around 1-year below average. Teacher 1 taught this group and is referred to as "Teacher 1B" in this study. During the first four observation sessions, there were two teacher-aids in the classroom, sitting with individual students. For the remainder of the sessions, there was one teacher-aid in the classroom. No data was collected on either teacher-aid. Observations were conducted during the second half of the school's first teaching block, immediately before morning tea.

Ethical Approval and Consent

Ethics approval was granted by the University of Waikato Division of Arts, Law, Psychology, and Social Sciences Human Research Ethics Committee (FS2020-10). Permission to conduct the study at the participating school was granted by one of the Deputy Principals, after consultation with the Principal, second Deputy Principal, and participating teachers. Before the study began, informed consent was provided by participating teachers (see Appendix A) and parental passive consent was obtained for all students in the participating groups (see Appendix C). After an ethics amendment was granted, parental consent was given for Nikau, before his results were reported (see Appendix B).

Materials

Data Collection

I developed separate data collection forms to record student and teacher behaviour. On-task student behaviour forms included; information on the name of the class and teacher being observed, operational definitions of on-task behaviour, with examples and non-examples; names of each student being observed, in relation to their assigned number; interval number; and number of the student to observe during each interval. There was also space to record; the date and time of observations; the name of the observer; any field notes relevant to the class; whether the observed student was engaging in on-task behaviour or not; any field notes relevant to the observed student during each interval; the total number of intervals each student was observed behaving appropriately; and the total number of intervals each student was present in the classroom during each session (see Appendix D). Teacher behaviour data collection forms were used to collect data on student disruptions and teacher praise statements, and included operational definitions of; praise and corrective statements; group and individual targets; and specific and general statements. Space was also available for recording; the date and time of the observation session; the class and teacher being

observed; the name of the observer; any general field notes relevant to the teacher; the interval number during which the behaviour occurred; the type of behaviour; any field notes relevant to each occurrence of behaviour; and the total number of each type of behaviour recorded during the session (see Appendix E).

A mobile device with a timing application uploaded onto it, which delivered an audio-recording and vibration at the end of each interval was used to time sessions. Headphones were used to decrease distractions within the classroom.

Procedural Integrity. A 13-item procedural integrity checklist was created to ensure each teacher received the same training on how to implement the tootling intervention (see Appendix F; Adapted from Lum, 2017). A second, 8-item, procedural integrity checklist was created to ensure the necessary steps were implemented by the teacher, to adequately inform and train the students (see Appendix G; Adapted from Lum, 2017).

Treatment Integrity. A daily treatment integrity checklist for teachers was produced to ensure adequate intervention implementation (see Appendix H; Adapted from Lum, 2017). A second (researcher) treatment integrity checklist was designed, which included items that could be observed directly, by the observers (see Appendix I).

Social Validity. On completion of the intervention, teachers completed a Modified Behaviour Intervention Rating Scale (BIRS; see Appendix M) to assess the social validity of tootling (Von Brock & Elliott, 1987). The BIRS is a 24-item questionnaire with a 6-point Likert scale, ranking responses from strongly disagree (1) to strongly agree (6), and measuring three factors: Acceptability, Effectiveness, and the Time of Effect (Von Brock & Elliott, 1987). According to Von Brock and Elliott (1987) these three factors produced coefficient alphas of .97, .92, and .87, respectively. Further research has found the BIRS to have high internal consistency ($\alpha = .92$) and good content and construct validity (Elliott & Treuting, 1991). The BIRS was modified for the purpose of this study to refer to past-tense

instead of future-tense, “children” instead of “child”, and “tootling instead of “intervention”. Modifications such as these have previously been found to have no impact on the psychometric properties of the instrument (Sheridan, Eagle, Cowan, & Mickelson, 2001). Example items include, “Tootling was an acceptable intervention for children’s problem behaviour”, “The intervention quickly improved the children’s behaviour”, and “Tootling produced a lasting improvement on children’s behaviour”.

Students with parental passive consent were invited to anonymously complete a modified Children’s Intervention Rating Profile (CIRP; see Appendix N) to assess student acceptability of the tootling intervention (Elliott, 1986). The CIRP is a seven-item questionnaire requiring students to rate their acceptance of an intervention on a 6-point Likert scale, ranging from strongly disagree (1) to strongly agree (6) (Elliott, 1986). The CIRP has high internal reliability, with a reported average coefficient alpha of .86 (Turco & Elliott, 1986), and has previously been used to measure student acceptability of tootling interventions (Lum et al., 2019; McHugh et al., 2016). The CIRP was modified for the purpose of this study, specifically identifying tootling as the intervention type and using smiley-faces, instead of numbers, to represent the Likert scale. Similar modifications have previously been found to maintain good internal consistency (Mitchell et al., 2015).

Tootling

I developed a script for teachers to use during the student tootle training, which specified what tootling was and how it would work in each classroom (see Appendix J). Slips of paper (13.5cm x 9cm) were provided for recording prosocial behaviour. In Groups 1 and 3, these *tootle cards* had space for students to write the name of the student observed and the behaviour performed. Alongside these tootle cards, Group 2 also had tootle cards which had space to write the observed student’s name, plus a list of five behaviours, which they could tick, to indicate which behaviour they had observed (see Appendix K). A tootle box (length:

22cm, height: 8.5cm, width: 15cm) was provided for placing written tootles, together with a progress chart with the numbers 1-100 and markers indicating the tootle target and class progress towards that target (see Appendix L).

Dependent Variables

The primary dependent variable during this study was student on-task behaviour. Secondary dependent variables were teacher corrective statements, which provided data on student disruptions, in a functional definition, and teacher praise statements. This teacher behaviour was further categorised by who it was targeted at (*individual* or *group*) and whether the specific behaviour was described (*specific*) or not (*general*). Daily number of tootles was also recorded to provide a measure of treatment integrity.

Student Behaviour

On-task behaviour. After consultation with the teachers and preliminary observations, on-task behaviour was operationally defined as a student; following teacher instructions and classroom rules; attending to (eyes oriented towards) appropriate materials, teacher, or peers, or eyes closed if appropriate to task, such as mindfulness; and participating in independent seatwork, or peer- or group activities. Independent seat work was defined as 1 or 2 knees or buttocks touching seat, with all four chair legs on the ground, using appropriate stationary or equipment in a way in which it was designed to be used. For example, if task was writing with a pencil, student was holding pencil using pencil grip, while attending to workbook, or with pencil touching paper. If the task involved using a Chromebook, fingers needed to be touching keys or within 5cm above keys.

Student disruptions. Student disruptions were measured functionally, by recording teacher corrective statements. Teacher corrective statements were defined as a vocal statement or gesture (such as shaking the head or raising the palm of a hand towards a student), in response to disruptive student behaviour, which functioned to stop or change that

behaviour. Corrective statements included requests to stop disruptive behaviour, such as "Please stop doing that" and "Don't throw darts", as well as verbal warnings, such as "This is your final warning" and "If you can't settle, you're going to have to move". Non-examples of corrective statements included chanting, "Tahi, rua" or "Shh, shh, sh, sh, sh" to gain the attention of all students. Corrective statements in response to academic tasks, such as, "No, that is not the correct answer" were also not included. A functional definition was chosen, instead of a topographical one, whereby only the form of behaviour is observed, because a functional definition provides a feasible way of collecting whole-class data and considers the effect of behaviour on the environment, regardless of its form. One example of a disruptive behaviour commonly described topographically, in previous research, is *out-of-seat* behaviour, defined as a student being out-of-seat without permission (Graham-Day et al., 2010; Lum et al., 2019; McHugh et al., 2016). However, a student may be out of their seat, without permission, without being disruptive. For example, to pick up a pencil, dropped onto the floor. Seeking permission for this may be more disruptive to the class than simply picking up the pencil, and the teacher may be more appreciative of this behaviour than permission seeking behaviour. A topographical definition would result in disruptive behaviour being scored if permission was not sought, whereas a functional definition would require a corrective statement from the teacher for the same score. Although a functional analysis is usually required to determine the true function of behaviour, in this study the function of behaviour was assumed if teacher behaviour ceased when student behaviour altered or ended. Corrective statements were chosen instead of reprimands, as used in previous studies (Elswick & Casey, 2011; Rubow et al., 2018), because a reprimand requires an indication of disapproval, which is difficult to define, behaviourally. For example, both teachers used the vocalisation, "shhhhh" to lower the noise level in their classrooms (which was different to "shh, shh, sh, sh, sh" or chanting "tahi, rua", used regardless of noise level). There was often

no indication of disapproval in this vocalisation, but the function was to reduce disruptive behaviour, making it important to capture.

Teacher Behaviour

Teacher Praise Statements. Teacher praise statements were initially defined using definitions from previous studies which have measured praise in response to behavioural interventions (Elswick & Casey, 2011; Rubow et al., 2018). These definitions were then modified during preliminary observations, to ensure the current study was comparable to previous research, whilst keeping definitions relevant to the current participants. Praise was defined as a positive vocal statement or gesture indicating approval of appropriate behaviour. For example, "Well done", "you're sitting beautifully", or "I love the way you stopped, looked and listened". Non-examples included giving out tokens with no explanation and praise in response to correct academic responses, such as "well done, that's the correct answer".

Breakdown of Teacher Behaviour. Teacher corrective statements and praise statements were broken down into subgroups, depending on whether they targeted a group or individual and whether the statement was specific or general.

Group. Group was defined as the teacher addressing two-or-more students, without vocalising specific names. For example, "You guys need to stop" or "Well done to the children who are still on task".

Individual. Individual was defined as the teacher orienting eyes towards and/or vocalising the name of the student they were addressing. More than one student could be named at the same time. For example, "Alice, John, and Fraser, please stop" or "Moana, well done for raising your hand".

Specific. Specific was defined as the teacher labelling a specific behaviour. For example, "I love how quietly you are sitting" or "Do not put tokens in your mouth". A

corrective statement was also recorded as specific if, instead of specifying the behaviour to stop, it specified a request to initiate a behaviour that was incompatible with a disruptive behaviour. For example, “Shhhhh” or “be quiet” does not specify the behaviour to be stopped but requests a behaviour that is incompatible with talking.

General. General was defined as the teacher not labelling the behaviour. For example, “Well done, Mila” or “Stop that please”.

Number of Tootles

The number of correctly written tootles (as defined in the teacher script) was recorded each day, by the teachers, as a measure of treatment integrity. A tootle was deemed correct if it included the name of the student observed and the prosocial behaviour performed by that student.

Data Collection

Observation times were agreed between myself and the teachers and were dependent on class timetable and availability. Observations were 36-minutes long for Groups 1 and 3, and 42-minutes long for Group 2. This difference in observation length was because of the number of students being observed in each group. Six students were observed from Groups 1 and 3, and seven students were observed from Group 2. Overall, each student was observed during 24 15-second intervals, across all groups. Each group was observed up to 3 times per week for 9 weeks, followed by a 7-week break (including a two-week school holiday break), before two maintenance observations were completed. During this time some observations were missed because of school closures, changes to class timetables, teacher or researcher illness, or teachers training students on the tootling procedure. Six, five, and five sessions were missed for groups 1, 2, and 3, respectively. Further, teacher behaviour was not observed once for Teacher 1A and twice for Teacher 2, as release teachers were in the classrooms in place of the usual teacher, during these sessions. On-task behaviour was recorded during

these sessions as an indication of generalisation of behaviour across teachers. Finally, Nikau was absent from one observation session. I (and the secondary observer, when present) was positioned in an unobtrusive location within the classroom, with a view of all students and the teacher, and within-group observations occurred at a consistent time each day, on the same days each week.

Observation Methods

Appropriate behaviour was recorded using a 15-second momentary time sampling technique, with an audio-recording indicating the end of each interval. Each student in each group was allocated a number prior to all observations and observed sequentially during each consecutive interval. That is, student 1 was observed during interval 1, student 2 during interval 2, student 3 during interval 3, etc. Once all students had been observed once, student 1 was observed again, followed by student 2, and so on. Nikau was student 7 and was observed using the same method. Momentary time sampling was chosen because it has been found to provide a more accurate estimate of actual behaviour, with fewer observer errors, compared to interval recording techniques (Alvero et al., 2008; Meany-Daboul et al., 2007). Fifteen-second intervals were chosen to allow time between intervals to record student and teacher behaviour without compromising accuracy (Gunter et al., 2003; Wirth et al., 2014). Although previous tootling studies observed whole-class behaviour using an interval-fixed method of observation (Dillon et al., 2019; Lum et al., 2019), a selection of students were observed during the current study because student movement was high, and accurately monitoring a selection of students was more manageable than whole-class observations. If students left the classroom for any reason and were no longer visible to the observers, each corresponding interval for that student was removed from the total number of intervals. Class-wide percentage of intervals engaged in on-task behaviour was estimated by dividing the total number of intervals of occurrence, across all students by the total number of

intervals in which the students were visible, multiplied by 100. Percentage of intervals engaged in on-task behaviour for Nikau were calculated in the same way, using only data collected on him.

All teacher behaviour (praise statements and corrective statements, as a measure of student disruptions) was recorded using event recording and each occurrence of behaviour was recorded on the data sheet, described above, along with the interval in which it occurred. This data recording system aligned with previous research measuring discrete instances of teacher behaviour (Elswick & Casey, 2011; Lannie & McCurdy, 2007). Rate of each behaviour per minute was calculated by totalling the number of responses during the entire observation duration, then dividing this number by the number of standard units of time in which the observations were conducted. For example, if a behaviour occurred a total of 30 times within a 36-minute observation period, the calculation would be: 30/36 = a rate of 0.83 occurrences per minute. No data on student disruptions for Nikau was collected. Data collection methods remained consistent throughout all phases and across all groups and participants.

Research Design

A multiple baseline across groups design, with follow-up phase, was used. Introduction of the intervention was staggered across groups to demonstrate that any change in behaviour was due to the intervention and not extraneous variables. An AB design, with follow-up phase was used for Nikau because results from his on-task behaviour were analysed separately once data collection was complete.

Phase-change decisions were made using visual analysis of level, trend, and variability of on-task behaviour. A minimum of five data points in each phase was needed before making any phase changes, as recommended to maintain confidence in data reliability (Kratochwill et al., 2010).

Procedure

Preliminary Observations

Prior to study commencement, 6 preliminary observations (2 for each group) were completed to; allow me to become familiar with student and teacher behaviour and accurately operationally define dependent variables; finalise observation techniques, with consideration of ease of use, accuracy, and reliability; and help normalise observer-presence, reducing participant reactivity. During this time, I determined that accurately capturing whole-class student disruption was not possible, using a topographical definition of behaviour, and hence used a functional definition hereafter. Results from these observations were not recorded. After the initial two recorded observations for Groups 1 and 3, Teacher 1 commented on student passive off-task behaviour during computer work. This involved the students sitting in front of computers with eyes oriented towards screens, but not actively engaged in work. After this, the definition of on-task behaviour was altered to include position of fingers on- or within 5cm above computer keys, to capture this behaviour more accurately. Data on on-task behaviour from these two initial observations were subsequently removed because dependent variables had changed. No amendment was needed for Group 2.

Baseline

Baseline data was collected on student on-task behaviour, disruptions, and teacher praise statements for a minimum of five sessions, prior to any training on intervention implementation. During this time, teachers were asked to conduct their classroom management strategies as normal.

Teacher Training

Once the decision to introduce the intervention to each group was made, I met with each teacher to discuss how the intervention would be implemented in their classroom. The script on how to train the students was provided to the teachers and discussed. Any questions

were answered, and the teachers added their own notes to the script to increase ease of use during student training. All tootling materials were provided to the teacher at this time.

Student Training

The teachers used the annotated script to explain the tootling procedure to the students and train them how to observe and record peer prosocial behaviour during class time. The teachers gave examples and non-examples of correct tootles and asked the students to give their own vocal and written examples. Appropriate praise and corrective feedback were provided by the teacher in response to the examples. Students were told where to post their tootles and when they could do this, and about the tootle target and group reward. Finally, students were asked to suggest ideas of group rewards. Rewards were then voted on and the most popular rewards were chosen for reaching the target.

Class-Wide Rewards. Rewards chosen by Group 1 were free time for reaching the first three targets of 60 tootles, followed by shared hot chips for reaching the fourth target of 90 tootles. Rewards chosen by Group 2 were an outside game, free time, free time, and a shared class lunch, for reaching a target of 40, 60, 70, and 70 tootles, respectively. Rewards chosen for Group 3 were 10-minutes free time on an electronic device at the end of class, each time they reached a target of 25 tootles (four times).

Tootling Intervention

Group 1 started their tootling intervention phase after the on-task data indicated a stable baseline. Group 2 began their tootling intervention phase after a stable trend was observed during Group 1's intervention. Group 3 began their tootling intervention phase after a stable trend was observed during Group 2's intervention. On the first day of the tootling intervention, after student training, the progress chart was added to the wall, in a location easily viewed by all students, and the tootle container was placed in a visible location. For Group 1 and Group 3 this was on an empty desk at the back of the classroom.

For Group 2 it was on a shelf under the whiteboard, at the front of the classroom. A selection of note cards was placed next to the tootle container. Group 1 students were also advised they could keep tootle cards in their desks if they preferred.

Fifteen minutes before the end of the school day (or 15 minutes before the end of the class for Group 3), teachers prompted students to place any final tootles in the container, selected three-to-five tootles, and silently read them, checking for accuracy, before reading them aloud to the class and praising students who the tootles were about. These praise statements were not included in the data collection because no observers were present during this time for Groups 1 and 2. Finally, the teacher counted the tootles in the container and added the number to the progress chart, along with any tootles previously added. Once the target was reached, the group reward was either issued immediately or the following day, depending on the reward and the time needed. For example, if the reward was 15-minutes free time, but there were only 5 minutes left in the school day, the reward was issued the following day. The progress chart was then reset back to zero and students began working towards the next goal.

Maintenance

During discussions with both teachers regarding fading of the intervention, both teachers communicated that they would prefer not to participate in a fading process consisting of systematic removal of tootling materials. Once all observations were complete for the intervention phase (which coincided with the end of term 3), teachers were told there would be break from observations and that they had the choice of whether to continue with tootling or not during term 4. Five-weeks after the final observations, I met with the teachers to provide a summary of the results. This was done in the form of a visual graph and a discussion. Seven-weeks after the final intervention observations, I returned to do two maintenance observations for each group (six in total). These observations were identical to

all previous observations, regardless of whether or not the teacher was continuing to use the tootling intervention.

Social Validity

Due to the reading skill level of some of the students in Group 2, the teacher read the CIRP questions aloud to the class and the students each rated the question by circling the appropriate smiley-face. Students who chose not to complete the CIRP were instructed to complete another activity during this time. Informed consent for Nikau was obtained after the CIRP had been completed by students in Group 2. Therefore, there is no individual data on Nikau's results of the CIRP.

Interobserver Reliability

Interobserver agreement (IOA) was measured to assess the accuracy and reliability of the observation data between myself and a secondary, trained observer. We independently and simultaneously recorded student and teacher behaviour, whilst sitting within proximity of each other but far enough apart so that neither could see what the other was recording. This ensured that we were able to observe the same behaviours in each classroom, especially as space was limited. The secondary observer was a Master's of Applied Psychology student who was trained on all observation and recording procedures, and operational definitions of target behaviour by myself. All IOA training occurred prior to any data collection. A minimum 80% IOA was required prior to- and during all data collection. If IOA scores fell below 80%, retraining of procedures and operational definitions occurred. Retraining was needed on three occasions for observations of on-task behaviour, after two IOA scores of 78% and one of 79%, after an IOA score of 79% for student disruptions, and after an IOA score of 75% for teacher praise statements. Table 1 shows the mean IOA scores for each type of behaviour across groups. Overall, mean percentage of IOA scores never fell below 80%.

Table 1

Mean percentage IOA Scores for each Type of Behaviour across Groups

Type of Behaviour		Mean % IOA Score (range)		
Student		Group 1	Group 2	Group 3
On-task	Baseline	97 (98 - 97)	99 (98 - 100)	91 (79 - 100)
	Tootling	89 (78 - 100)	90 (83 - 95)	80 (78 - 84)
	Follow-up	88 (no range)	84 (no range)	92 (no range)
	Across Phases	91	91	88
Disruptions	Baseline	84 (80 - 88)	85 (81 - 91)	88 (82 - 100)
	Tootling	91 (85 - 100)	84 (84-85)	81 (79 - 83)
	Follow-up	88 (no range)	89 (no range)	90 (no range)
	Across Phases	88	86	86
Teacher		Teacher 1A	Teacher 2	Teacher 1B
Praise Statements	Baseline	100 (no range)	85 (82 - 88)	85 (75 - 100)
	Tootling	100 (no range)	86 (84 - 88)	100 (no range)
	Follow-up	100 (no range)	83 (no range)	100 (no range)
	Across Phases	100	85	95

IOA was obtained for 17 (25.76%; range: 21.43% - 40%) of the study's 66 sessions, which met the recommendation by previous researchers on IOA of at least 20% across phases and groups (Kennedy, 2005). Two methods of calculating IOA for on-task behaviour were used, depending on the percentage of occurrence, to ensure the most conservative methods were always used. When the frequency of behaviour was above 70%, the *unscored-interval IOA* technique was used, whereby intervals with non-occurrence of behaviour recorded by either, or both, observers were used in the calculation (Cooper et al., 2014). Of these intervals, IOA was calculated by dividing the total number of agreements by the total number of agreements plus disagreements and multiplying by 100. When behaviour fell between 30% and 70%, the *interval-by-interval IOA* technique was used by calculating the total number of intervals agreed, divided by the total number of intervals disagreed plus the total number of intervals agreed, multiplied by 100 (Cooper et al., 2014). For Group 1, IOA was measured during 29% of the baseline phase, 21% of the tootling phase, and 50% of the maintenance phase.

For Group 2, IOA was measured during 25% of the baseline phase, 30% of the tootling phase,and 50% of the maintenance phase. For Group 3, IOA was measured during 24% of the baseline phase, 40% of the tootling phase, and 50% of the maintenance phase.

To calculate IOA scores for student disruptions and teacher praise statements, intervals with 100% agreement were divided by the total number of intervals whereby occurrence of behaviour was recorded by either, or both, observers, multiplied by 100. For Teacher 1A, IOA was measured during 29% of the baseline phase, 23% of tootling phase, and 50% of the maintenance phase. For Teacher 2, IOA was measured during 25% of the baseline phase, 25% of the tootling phase, and 50% of the maintenance phase. For Teacher 1B, IOA was measured during 24% of the baseline phase, 40% of the tootling phase, and 50% of the maintenance phase.

Teachers counted and recorded the number of tootles collected each day, before giving them to me. I then counted the tootles again for agreement purposes. Agreement for number of tootles across groups was 100%. Treatment integrity IOA was also obtained during 28% of observations and was 100%.

Data Analysis

Visual Analysis

All data were entered into Microsoft Excel after each session and graphs were produced using this software. Visual analysis was used to evaluate level, trend, and variability of the data within- and between-conditions (Lane & Gast, 2013), allowing decisions to be made regarding appropriate timing of new phase introduction, as well as determining any functional relationships between the intervention and the dependent variables (Kratochwill et al., 2010).

Effect Size Calculations

Treatment effects were also evaluated using Tau-U effect size calculations. In previous tootling literature, the most frequently used calculations of effect size were non-overlap of all pairs (NAP; Parker & Vannest, 2009), used by Lambert et al. (2015), Lipscombe et al. (2018), Lum et al. (2018), and McHugh et al. (2016), and tau for nonoverlap with baseline control (Tau-U; Parker et al., 2011), used by Kirkpatrick et al. (2019), Lum et al. (2019), and Lum et al. (2017). NAP, an index of data overlap between phases; calculates overlap of paired data points between baseline and treatment phases; was designed to overcome limitations of previous indices of nonoverlapping data; and was found to be strongly associated with the commonly published effect size, R^2 (Parker & Vannest, 2009). Tau-U uses similar calculation methods to NAP, but controls for positive baseline trend, shows no artificial ceiling effects, and has been found to be more conservative than NAP (Parker et al., 2011). Considering the above, the more conservative Tau-U effect size calculation was used in the current study, which still allowed comparisons to be made with previous findings. In accordance with Vannest and Ninci (2015), effect sizes were interpreted as .00 to .20 indicating a small change, .20 to .60 indicating a moderate change, .60 to .80 indicating a large change, and above .80 indicating a very large change. Baseline data were analysed for any significant trends, with a trend level score above .4 in the appropriate direction requiring a trend correction. No trend corrections were required for any behaviours.

Results

Measures of Integrity

Procedural integrity

Procedural integrity scores were calculated by dividing the number of steps completed by the total number of steps identified on the integrity form and multiplying by 100. Integrity scores during each teacher and student training session were 100%.

Treatment integrity

Mean teacher-rated treatment integrity was 95% for teacher 1A, 100% for teacher 2, and 94% for teacher 1B. Mean observer-rated treatment integrity was 96% for teacher 1A, 100% for Teacher 2, and 75% for teacher 1B. Teacher 1A reported not completing the final steps, once tootling had finished for the day, on two occasions. These steps were, step 5, *Read several tootles and provide praise/corrective feedback,* step 6, *Count tootles and provide reward if target achieved,* and step 7, *Update progress chart*. Teacher 1B reported not completing the same steps as above, on one occasion. During my own observations, I recorded the feedback chart not being updated for Group 1 (Teacher 1A) on two occasions, alongside the final three steps listed above not being completed on three occasions for Group 3 (Teacher 1B). Although Teacher 1A did not report ever not providing the reward once the tootle target was met, according to field observations, the teacher was reminded by students on 3 occasions that they had not received their reward (free time) for reaching their target.

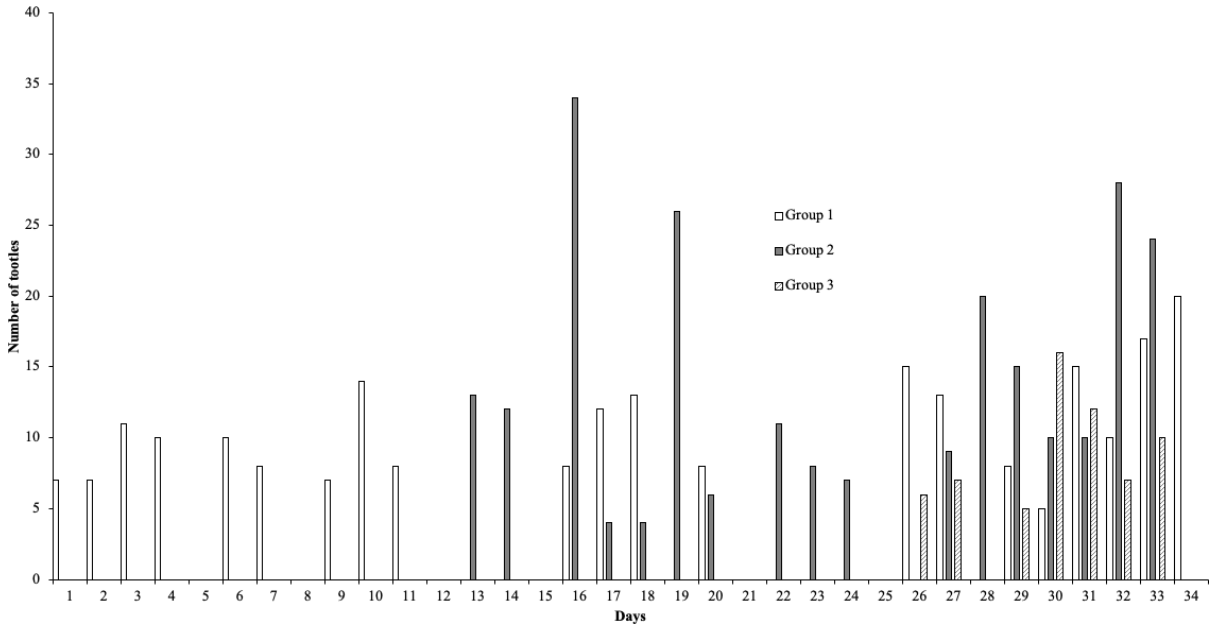
Number of Tootles

Tootles were counted by the teacher at the end of each day and recorded for integrity purposes. Figure 1 shows the number of tootles collected each day in each classroom during the intervention phase. This data represents every day of the intervention phase, across groups, not the days when observations occurred, which was only three days per week. The teacher reported the number of tootles for Group 1 on 21 (62%) of the 34 days that tootling

was in place. The number of tootles for Group 2 was reported on 17 (77%) of the 22 days that tootling was present in the classroom. Finally, the number of days tootles were reported for Group 3 during the intervention was 7 (78%) out of 9. The mean number of daily tootles collected throughout the intervention for Groups 1, 2, and 3 was 10.27 (range = 0 - 20), 14.18 (range = 4 - 34), and 9.00 (range = 5-16), respectively.

Figure 1

Tootle Count as Reported by Teachers across Groups during Intervention Implementation



Student Behaviour

Student on-task behaviour was the primary dependent variable and student disruptions, measured using teacher corrective statements, was a secondary dependent variable. Figure 2 shows the aggregated percentage of on-task behaviour and rate of disruptions per minute for each group across phases. The first two data points for on-task behaviour of Groups 1 and 3 are missing due to amendments being made to the operational definition after comments from the teacher regarding much passive off-task behaviour. Table 2 shows the mean, range, and Tau-U scores of on-task behaviour and disruptions across the three groups. Effect sizes were calculated between the baseline and tootling phases (see Table

2). A release teacher was in the classroom during three observation sessions: session number 10 for Group 1 and session numbers 17 and 19 for Group 2. Data on student on-task behaviour was still collected during these sessions, but not included in the mean, range, or effect size calculations. Teacher behaviour was not observed during these sessions; therefore, no data was collected on student disruptions, which required the presence of the participating teachers. These data are included in Figure 2 for generalisation purposes and are depicted as open squares.

Figure 2

Percentage of Student On-Task Behaviour and Rate of Student Disruptions across Sessions

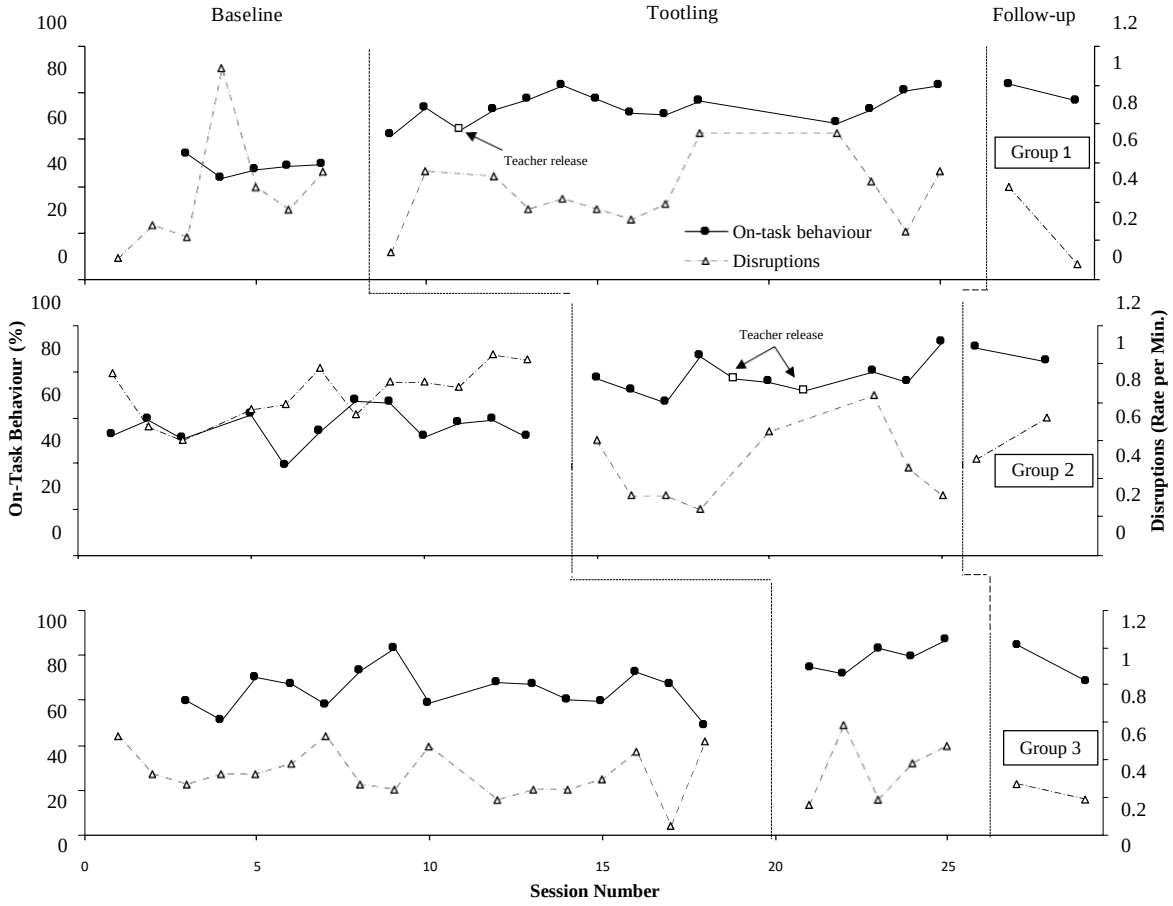


Table 2

Means, Ranges, and Tau-U Scores for On-Task behaviour and Disruptions during each Phase across Groups.

Behaviour	Phase	Group 1		Group 2		Group 3	
		Mean (range)	Tau-U	Mean (range)	Tau-U	Mean (range)	Tau-U
		%		%		%	
On-Task	Baseline	48 (44 - 53)		56 (39 - 67)		66 (49 - 83)	
	Tootling	74 (62 - 83)	1**	78 (66 - 93)	.96**	79 (71 - 87)	.84*
	Follow-up	80 (76 - 84)		87 (84 - 90)		77 (68 - 85)	
		rate/min		rate/min		rate/min	
Disruptive	Baseline	0.44 (0.11 - 1.08)		0.85 (0.60 - 1.05)		0.34 (0.06 - 0.53)	
	Tootling	0.45 (0.14 - 0.75)	.19	0.46 (0.24 - 0.83)	-.86**	0.36 (0.17 - 0.58)	.05
	Follow-up	0.28 (0.08 - 0.47)		0.61 (0.50 - 0.71)		0.24 (0.19 - 0.28)	

** Very large effect size
* Large effect size

Group 1

On-task Behaviour. During baseline, mean percentage of on-task behaviour for students in Group 1 (see Figure 2, top panel) was 48% (range = 44% - 53%; Table 2), with stable responding, before tootling was introduced. Behaviour immediately increased from 49% at the end of baseline to 62% at start of the intervention. During the tootling phase, behaviour was slightly variable with a generally increasing trend and a mean of 75% (range = 62% - 83%). The follow-up data (*M* = 80%; range = 76% - 84%) shows little change in on-task behaviour during the first follow-up session (84%), compared to the last session in the tootling phase (83%), followed by a decrease in behaviour during the final session (76%). Overall, using the standards set by Vannest and Ninci (2015), effect size for increasing on-

task behaviour in Group 1 was very large (1; see Table 2), indicating a very high level of non-overlap between baseline and intervention pairs and a highly effective intervention.

During the teacher-release session for Group 1, on-task behaviour was 64%. This was a decrease from the previous session (74%) but an increase from the first session in the tootling phase (62%). During the session following that with the release teacher, on-task behaviour increased to 73%.

Disruptions. For Group 1, during baseline, mean rate of student disruptions was 0.44 per minute (range = 0.11 – 1.08), with an increasing trend and some variability. During the intervention phase, rate of disruptions immediately dropped from 0.56 per minute, during the final baseline session, to 0.14 during the first intervention session. Mean rate of disruptions during the intervention phase was 0.47 per minute (range = 0.14 – 0.75) with high variability. The mean dropped to 0.28 (range = 0.08 - 0.19) disruptions during the two follow-up sessions. Overall, there was a small negative effect size (-.19), indicating a low level of non-overlap between baseline and intervention pairs and little effect of tootling on decreasing student disruptions.

Group 2

On-Task Behaviour. Mean percentage of on-task behaviour during baseline, for students in Group 2 (see Figure 2, middle panel), was 56% (range = 39% - 67%; Table 2), with moderate variability. During the intervention, behaviour increased from 51% at the end of baseline to 77% at the start of tootling. After this, there was a mostly increasing trend, with some variability and a mean of 78% (range = 66% - 93%). On-task behaviour remained high during follow-up with a mean of 87 % (range = 84% - 90%). Overall, Tau-U calculations resulted in a very large effect size (.96) for increasing student on-task behaviour in Group 2, indicating a high level of non-overlap between pairs.

During the 5th session for Group 2, when the release teacher was present, percentage of on-task behaviour was 77%, which decreased from 86% in the previous session, with the usual teacher, followed by 76% in the following session with the usual teacher. During the 7th session, whereby the release teacher was present again, on-task behaviour remained stable (72%), before increasing (80%) once the usual teacher returned.

According to field observations, Teacher 2 implemented a change to the tootling procedure during the second and third days of the tootling phase. This adaptation involved the teacher instructing students to write more meaningful tootles, for which they could earn double points on the progress chart. The adaptation continued over 2 days before the double-point incentive was removed. This data equates to days 17 and 18 in Figure 1, which shows the least number of tootles was collected for Group 2 on these days.

Disruptions. During baseline, mean rate of student disruptions for Group 2 was 0.85 disruptions per minute (range = 0.60 – 1.05). Behaviour was variable, with a slightly positive trend. Disruptions immediately decreased during the intervention, from 1.02 at the end of baseline to 0.60 at the start of the intervention. Behaviour continued to be variable during the tootling phase (M = 0.46; range = 0.24 - 0.83) and increased during follow-up (M = 0.61; range = 0.50 - 0.71). Tau-U calculations resulted in a very large effect (.87) for decreasing rate of student disruptions for Group 2.

Group 3

On-Task Behaviour. For Group 3 (see Figure 2, bottom panel), during baseline, mean percentage of on-task behaviour was 66% (range = 49% - 83%; Table 2), with high variability. Once the intervention was implemented, behaviour immediately increased from 49% at the end of baseline to 75% at the start of the intervention. During the tootling phase, there was a more stable, increasing trend in behaviour and a mean of 79% (range = 71– 87%). On-task behaviour remained high (85%; M = 77%) during the first follow-up session, then

decreased to 68% during the final session. A very large effect size (.84) for increasing on-task behaviour was found, indicating a high level of non-overlap between baseline and intervention pairs.

Disruptions. For Group 3 during baseline, mean rate of student disruptions per minute was 0.34 (range = 0.06 – 0.53), with some variability. Rate of behaviour immediately decreased from 0.50 during the final baseline session to 0.17 during the first intervention session, with a mean of 0.36 (range = 0.17 – 0.58) and some variability, initially during this phase. Rate of behaviour dropped from 0.47 at the end of the tootling phase to 0.27 during the first follow-up session, then remained stable during the final follow-up session (0.19). A small effect (.05) for increasing student disruptions was found, between baseline and intervention phases, for Group 3.

Individual Student from Group 2

Only on-task behaviour across the three phases was measured for Nikau, because his individual data was not originally part of the study, which can be seen in Figure 3. During baseline, mean percentage of on-task behaviour was 52% (range = 29% - 71%), with some variability. At the start of the intervention, on-task behaviour immediately increased to 82 %, from 52% at the end of baseline. During the tootling phase, variability continued, but mean percentage of on-task behaviour increased to 79% (range = 61% - 92%). During the follow-up phase (*M* = 83%; range = 75% - 92%) an increase in on-task behaviour during the first follow-up session (92%), compared to the last session in the tootling phase (88%), followed by a decrease in behaviour (75%), during the final session was observed. A very large effect size (.95; See Table 2) for increasing Nikau's on-task behaviour was found from Tau-U calculations, indicating a high level of non-overlap between baseline and intervention pairs, and a very successful intervention for Nikau.

During a meeting with Teacher 2, 10-weeks following the end of the tootling intervention, Teacher 2 explained that Nikau had recently met with his Resource Teacher: Learning and Behaviour, to discuss Nikau's progress throughout the year. In this meeting Nikau had been asked what supports and strategies he felt had helped him at school. When discussing strategies to help him manage his emotions, Nikau commented, "Tootling has helped me". This personal response aligns with the results from the current study, indicating that tootling had a positive effect on Nikau's behaviour.

Figure 3

Percentage of On-Task Behaviour of the Target Student across Phases

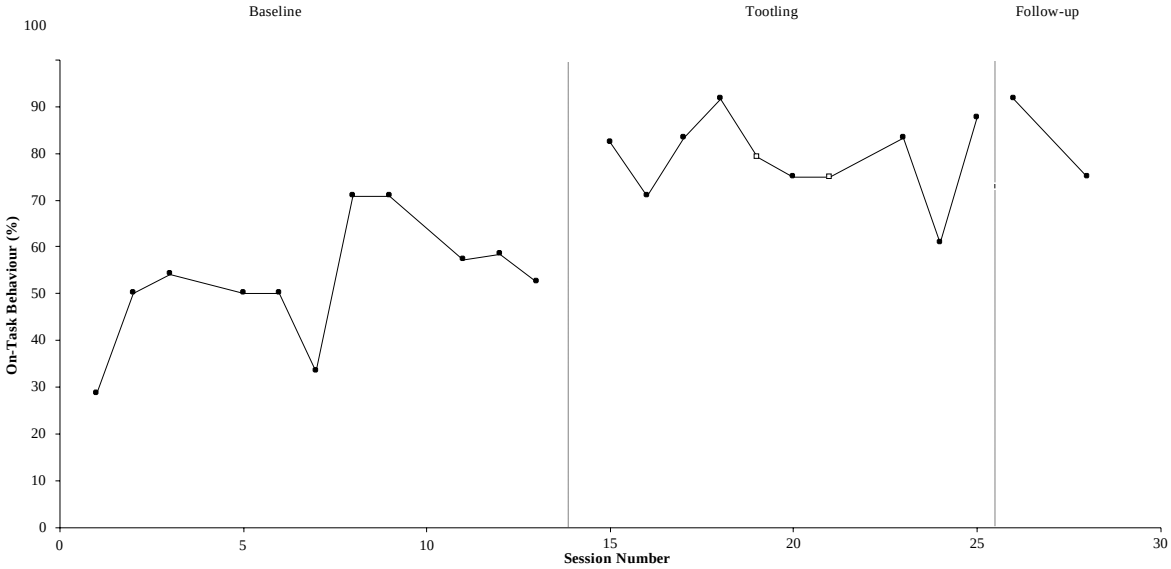


Table 3

Means, Ranges, and Effect Sizes for On-Task Behaviour of Target Student

Phase	Mean (range)	Tau-U
Baseline	52 (29 - 71)	
Tootling	79 (61 - 92)	.95*
Follow-up	83 (75 - 92)	

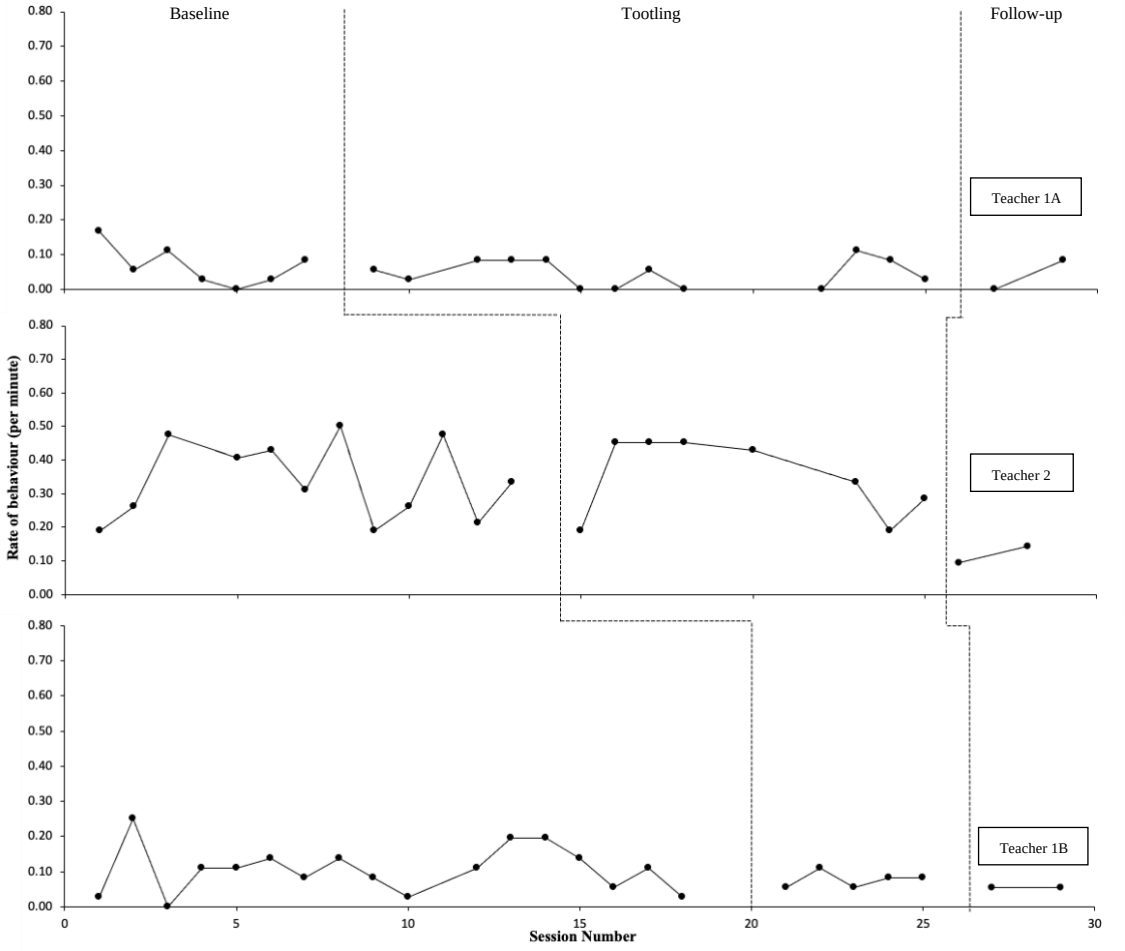
*Very large effect size

Teacher Praise Statements

Teacher praise statements was another secondary dependent variable and results can be seen in Figure 4.

Figure 4

Rate of Teacher Praise Statements for each Teacher across Phases



Teacher 1A

Mean rate of total teacher praise statements for teacher 1A (see Figure 4, top panel), during baseline was 0.07 per minute (range = 0.00 - 0.17) with a mostly decreasing trend and little variability (See Figure 4). After intervention implementation, rates of praise remained stable, with a rate of 0.06 per minute during the first intervention session, following a rate of 0.08 per minute during the final session of the baseline phase. The mean rate of praise statements during the intervention phase reduced to 0.05 per minute (range = 0.00 - 0.11) and was stable throughout the phase. During follow-up, rates of praise were 0.00 and 0.08, respectively. A small effect size (-.20) in decreasing Teacher 1A praise statements during the

intervention phase, compared to baseline was found, indicating a counter-therapeutic effect of the intervention on praise statements.

Teacher 2

Mean rate of praise statements for Teacher 2 during baseline was 0.34 per minute (range = 0.19 - 0.50), with some variability (see Figure 4, middle panel). During the intervention phase, praise statements decreased from a rate of 0.33 per minute at the end of baseline to 0.19 at the start of the intervention. Mean rate of praise statements during the intervention was 0.35 (range = 0.19 - 0.45) with some variability. Mean rates of behaviour during sessions in the follow-up phase (0.10 and 0.14) were lower than any rates in previous sessions. A small effect (.02) for increasing praise statements during the intervention, compared to baseline, was found for Teacher 2.

Teacher 1B

Mean rate of praise statements for Teacher 1B during baseline was 0.11 (range = 0 - 0.25) with some variability (see Figure 4, bottom panel). There was little change in behaviour as soon as the intervention was introduced and rates went from 0.03 per minute, at the end of baseline, to 0.06 per minute, at the start of the intervention. Behaviour was stable during the intervention phase and had a mean rate of 0.08 (range = 0.06 - 0.11). Stability continued into the follow-up phase with praise occurring at a rate of 0.06 statements per minute across both sessions. Effect size calculations indicated a counter-therapeutic effect, with a moderate effect (-.29) for decreasing praise statements.

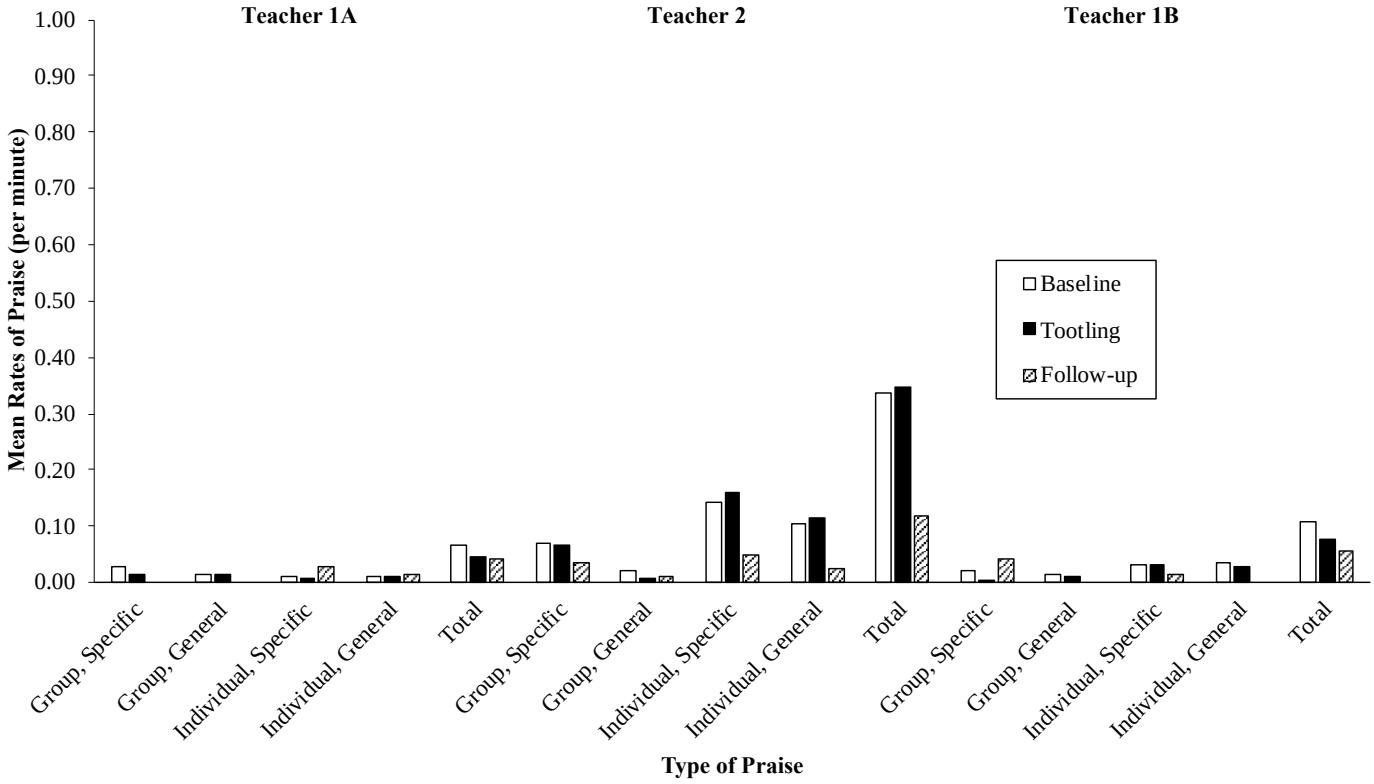
Types of Teacher Praise

Praise was further split into four groups, depending on the type (specific or general) and who was the target (individual or group). Figure 5 shows the mean rates of the different types of praise for each teacher across phases.

Teacher 1A. The rate of all types of praise, across all phases, for teacher 1A was low. During baseline, mean rate of group specific praise was 0.03 per minute, which decreased to 0.01 during the intervention and 0.00 during follow-up. Mean rate of group general praise was 0.02 during baseline, 0.01during intervention, and 0.00 during follow-up. Mean rates of individual specific praise began at 0.01 during baseline and remained consistent during tootling but increased to 0.03 during follow-up. Rates of individual general praise remained consistent across all phases (0.01).

Figure 5

Mean Rate of Group, Individual, General, and Specific Teacher Praise Statements during each Phase across Teachers



Teacher 2. Overall, Teacher 2 delivered more individual praise than group praise and more specific praise than general praise. During baseline and tootling phases, mean rate of group, specific praise was 0.07 per minute, which decreased to 0.04 during follow-up sessions. Mean rate of group general praise was 0.02 during baseline, then 0.01 during intervention, and follow-up. Individual, specific praise increased from a mean rate of 0.14

during baseline, to 0.16 during tootling, before decreasing to 0.05 during follow-up. Individual, general praise followed a similar pattern, with a mean rate of 0.10, 0.11, and 0.02 during baseline, tootling, and follow-up, respectively.

Teacher 1B. Overall, rates of all types of praise delivered by Teacher 1B were low. During baseline, mean rate of group, specific praise was 0.02 per minute, which decreased to 0.01 during the intervention and increased to 0.04 during follow-up. Group general praise was delivered at a mean rate of 0.01 during baseline and tootling, then 0.00 during follow-up. Mean rates of individual, specific praise was 0.03 during baseline and remained consistent during tootling, then decreased to 0.01 during follow-up. Rates of individual, general praise were delivered at a rate of 0.04 during baseline, then decreased to 0.03 during tootling and 0.00 during follow-up.

Teacher Corrections

Measuring student disruptions as a function of teacher corrections allowed this variable to be split in the same way as teacher praise statements, depending on the type (specific or general) and who was the target (individual or group). Figure 6 shows the mean rates of the different types of corrections for each teacher across phases.

Teacher 1A. During baseline, for Teacher 1A mean rate of group, specific corrections was 0.14 per minute, which decreased to 0.12 during the intervention and 0.08 during follow-up. For group, general corrections, mean rate was 0.04 corrections per minute during baseline and tootling, with a zero rate at follow-up. Mean rate of individual, specific corrections was 0.23 during baseline, 0.20 during tootling, and 0.17 during follow-up. Mean rate of individual, general corrections was 0.03 corrections per minute during baseline, then increased to 0.10 during the intervention, before returning to 0.03 at follow-up.

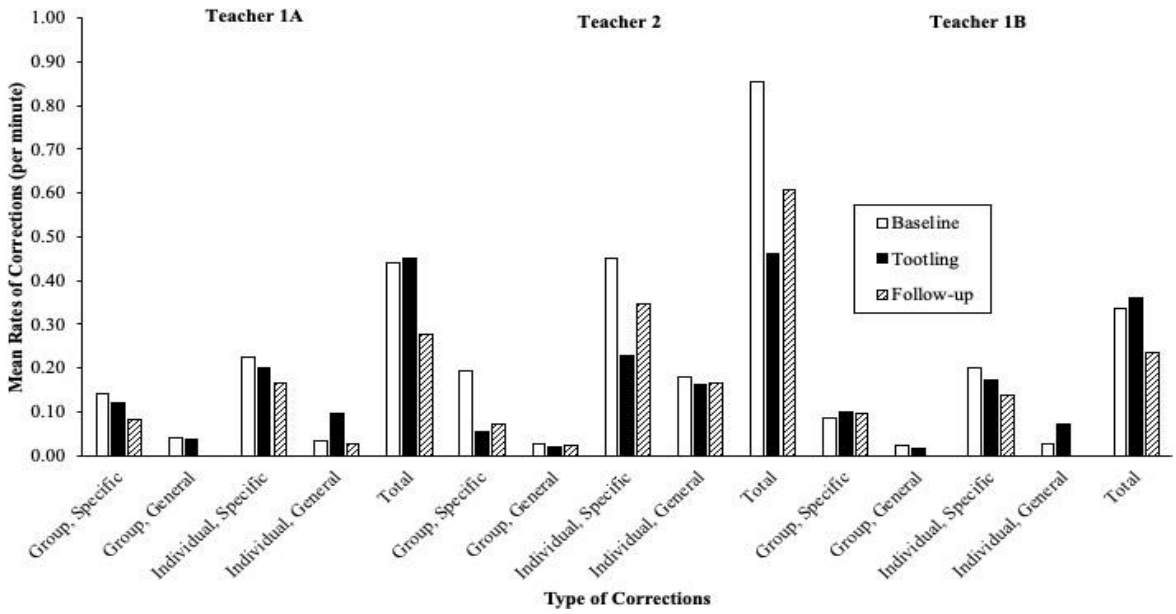
Teacher 2. For Teacher 2, the mean rate of group, specific corrections was 0.19 during baseline, then decreased to 0.05 during tootling, and 0.07 during follow-up. General

corrections remained consistent throughout the phases with a mean rate of 0.02 per minute during baseline and 0.03 during tootling and follow-up. Mean rate of individual, specific corrective statements was 0.45 during baseline, 0.23 during tootling, and 0.35 during follow-up. Mean rate of individual, general corrections was similar across phases, at 0.18 during baseline, 0.16 during tootling, and 0.17 during follow-up.

Teacher 1B. The mean rate of group, specific corrections was 0.08 per minute during baseline, then increased to 0.10 per minute during tootling and intervention phases. Group, general corrections was consistently low across phases, with a mean rate of 0.02 during baseline and intervention, then 0.00 at follow-up. For individual, specific corrective statements, the mean rate per minute decreased from 0.20 during baseline, to 0.17 during tootling, and 0.14 at follow-up. Mean rate of individual, general corrections was 0.03 during baseline, 0.07 during tootling, and 0.00 during follow-up.

Figure 6

Mean Rate of Group, Individual, General, and Specific Teacher Correction Statements during each Phase across Teachers



Praise to Correction Ratios

Again, using teacher corrective statements to define student disruptions allowed praise to correction ratios to be calculated (See Table 4). For Teacher 1A, ratios were 1:8, 1:8, and 1:5 during baseline, intervention, and follow-up phases. Teacher 2 had the highest ratios of 2:5 during baseline, 3:4 during tootling, which was the highest ratio of all teachers, across all phases, and 2:11 during follow-up phases. Ratios for Teacher 1B were1:4 during baseline, 1:7 during the intervention, and 1:4 during follow-up.

Table 4

Praise to Correction Ratios for all Teachers across all Phases

Teacher	Phase	Praise : Correction Ratio
1A	Baseline	1:8
	Tootling	1:8
	Follow-up	1:5
2	Baseline	2:5
	Tootling	3:4
	Follow-up	2:11
1B	Baseline	1:4
	Tootling	1:7
	Follow-up	1:4

Social Validity

To assess the social validity of tootling in each classroom, each participating teacher completed a BIRS (Elliott & Treuting, 1991), and those students who chose to (84%), completed a CIRP (Turco & Elliott, 1986), after the intervention phase was complete.

Teacher ratings

Each of the 24-items on the BIRS had scores ranging from 1 to 6, with higher scores indicating higher agreement. There was one exception to this, which required reverse scoring: Item 8 stated, “Tootling resulted in negative side-effects for some children”. Teacher 1 did not answer item 21, “Using tootling not only improved the children’s behaviour in the classroom, but also in other settings (e.g., other classrooms, home)”, for either of the BIRS

she completed, anecdotally reporting that this was “unknown” to her. Therefore, this item was removed from the calculations for Teacher 1A and Teacher 1B but remained for Teacher 2. Teacher 1A’s total score out of 144 was 106.5, Teacher 2’s score was 119, and Teacher 1B’s score was 109. Overall mean per item for Teacher 1A, Teacher 2, and Teacher 1B was 4.65 (range = 3.00 - 6.00), 4.96 (range = 3.00 - 6.00), and 4.74 (range = 3.00 - 6.00), respectively. Table 5 shows the mean rating, out of a maximum score of 6, for each teacher across the three factors measured by the BIRS: Acceptability, Effectiveness, and Time of Effect. The overall mean on the Acceptability factor was 5.10 (range = 4.83 - 5.47), on the Effectiveness factor was 4.11 (range = 4.00 - 4.17), and on the Time of Effect factor was 4.50, with all teachers scoring the same on this factor. In other words, overall, the teachers agreed that tootling was acceptable and slightly agreed that it was effective, with a quick rate of improvement in behaviour.

Table 5

Mean Scores across Acceptability, Effectiveness, and Time of Effect Factors, as Measured by the BIRS.

Factor	Teacher			Overall mean across factors
	1A	2	1B	
Acceptability	4.87	5.47	5.00	5.10
Effectiveness	4.17	4.00	4.17	4.11
Time of Effect	4.50	4.50	4.50	4.50
Overall Mean	4.65	4.96	4.74	4.78

Student ratings

Each of the seven items on the CIRP had scores ranging from 1 to 6 (represented by a collection of sad- and smiley-face-pictures), with higher scores (smiley-faces) indicating higher agreement. There were three exceptions to this, which required reverse scoring: Item number 2, “Tootling was too hard on me”, item number 3, “Tootling caused problems with my friends”, and item number 4, “There are better ways to handle problem behaviour than

tootling". Twenty-six (84%) of the 31 students in Group 1, 15 (71%) of the 21 students in Group 2, and 24 (96%) of the 25 students in Group 3 completed the CIRP. Mean item-scores, out of 6, were calculated, along with the overall item means (see Table 6). The overall mean item-score for Group 1 was 4.72 (range = 4.19 - 5.58), for Group 2 was 4.94 (range = 3.25 - 5.87), and for Group 3 was 4.53 (range = 3.48 - 5.42). Across all three groups, the overall mean item-score was 4.73 (range = 3.63 - 5.52).

Table 6

Mean Item Scores Within- and Between-Groups for the CIRP

Group	Mean item-score							Overall item mean
	1	2	3	4	5	6	7	
1	5.27	4.35	5.58	4.19	4.63	4.72	4.31	4.72
2	5.87	3.91	5.15	3.25	5.43	5.40	5.60	4.94
3	5.42	3.91	5.15	3.25	5.43	5.40	5.60	4.94
Mean	5.52	4.11	5.20	3.64	5.01	4.92	4.72	4.73

The individual mean score for each participating student within each class was calculated by adding the completed scores together and dividing by the number of items completed. These means were then totalled together and divided by the number of completed CIRP questionnaires, then multiplied by the number of questions in the CIRP (7) to give a mean total score out of 42 (see Table 7). In Group 1, two students did not complete item number 3, "Tootling caused problems with my friends" and one student did not complete item number 6, "I liked tootling". In Group 2, four students did not complete item number 2, "Tootling was too hard on me", two students did not complete item number 3, three students did not complete item 4, "There are better ways to handle problem behaviour than tootling", one student did not complete item number 5, "Tootling would help other children too", five students did not complete item number 6, and six students did not complete item number 7, "I think tootling would help me to do better in school". These missing items were removed from all calculations, which were then adjusted appropriately to ensure a true representation of the

data. For example, if one student rated all items 6/6 except two items, which was left blank, then the individual mean score for that student would be 5 (completed items) multiplied by (a score of) 6, divided by 5 (completed items). The mean total score, for Group 1 was 32.89 (range = 16 - 42), for Group 2 was 35.16 (range = 12 - 40), and for Group 3 was 31.91 (range = 24 - 42). Across all groups, the combined mean total score was 33.32. These mean total scores indicate that, on average, students found the tootling intervention to be acceptable, according to the criteria of 24.5 or above suggested by Turco and Elliott (1986). Of the students who completed all items, one from Group 1 (total score = 16) and one from Group 3 (total score = 24) did not find the intervention acceptable according to the above criteria.

Notes taken from field observations after all students had completed the CIRP indicate that numerous students in Group 2 did not understand some of the questions in the CIRP. Teacher 2 suggested that the reverse scoring used for some of the questions was confusing for the students and they were unsure how to respond appropriately.

Table 7

Mean Total Scores from the CIRP across Groups

Group	Mean Total Score
1	32.89
2	35.16
3	31.91
Combined Mean Score	33.32

Discussion

A shift in focus from punitive techniques to function-based interventions, based on ABA principles, means schools now have access to tools needed to promote a positive behaviour approach to classroom management. Teachers can feel overwhelmed with multiple competing classroom stimuli, making application of some behaviour management strategies difficult. Peer-mediated interventions can improve numerous classroom skills and behaviours. Tootling, alongside an interdependent group contingency and public display of progress is a peer-mediated intervention designed to teach students how to recognise and report peers' prosocial behaviour. Previous researchers, who have implemented tootling into various classrooms, have found promising results related to student behaviour, but none have measured the effects of tootling on teacher behaviour. In the current study, a single-subject multiple baseline design, with follow-up phase, was implemented in two, year 5/6 classrooms and one, year 1 classroom to investigate the effects of a tootling intervention on student and teacher behaviour.

Research questions

Question 1

The first research question in this study addressed whether a tootling intervention, implemented in year 1 and year 5/6 primary-school classrooms, would increase on-task behaviour, class-wide. Visual analysis of the results from the multiple baseline design shows a sudden increase in on-task behaviour, once tootling was introduced, which remained throughout the phase, compared to baseline, for all three groups. These results demonstrate a functional relationship between tootling and on-task behaviour and align with previous research on tootling, whereby researchers found increases in academically engaged behaviour during tootling phases, compared to baseline and withdrawal phases (Cihak et al., 2009; Dillon et al., 2019; Kirkpatrick et al., 2019; Lambert et al., 2015; Lipscomb et al., 2018; Lum

et al., 2019; Lum et al., 2017; McHugh et al., 2016). Although no previous studies have investigated the effects of tootling on on-task behaviour of students with a below average mathematics skill level, results from Group 3 indicate that tootling can also be beneficial for students with some skill deficits. Very large effect sizes for increasing on-task behaviour across groups were found in the current study. Similar results were found by Lum et al. (2019), who also used Tau-U to calculate effect size, and Dillon et al. (2019), who used Tau-U ranges (Vannest & Ninci, 2015) to describe raw scores from baseline corrected Tau-U calculations. Lum et al. (2017) reported weak to moderate effect sizes on academically engaged behaviour but used more conservative ranges from NAP to describe their Tau-U calculations, due to a lack of guidelines for Tau-U at the time of their research. However, raw Tau-U scores, calculated by Lum et al. (2017) were similar to the scores in the current study, indicating similar effects were found.

During two follow-up sessions, seven weeks after the tootling phase ended (including a 2-week school break), levels of on-task behaviour were maintained across groups during the first session. All on-task behaviour decreased during the second follow-up session, with Group 1 and 2 data remaining within the range of the tootling phase and Group 3 data dropping to baseline levels. Teacher 1A (Group 1) stopped using tootling once the intervention phase ended, Teacher 2 (Group 2) continued to use tootling on most days between the intervention and follow-up phases, and Teacher 1B (Group 3) continued tootling for 2-weeks of the new term, following the intervention phase, before stopping (3-weeks prior to follow-up). These results are mostly consistent with previous findings of maintained behaviour changes during follow-up trials, when teachers chose to continue to use tootling in their classrooms (Lambert et al., 2015) and a lack of maintenance when teachers did not (Lum et al., 2017). The teachers in the Lambert et al. (2015) study continued tootling in their classrooms during follow-up trials, as did Teacher 2, of Group 2, in the current study. Results

during follow-up for both studies were similar. One difference was that Teacher 2, in the current study, implemented a natural fading procedure, by only tootling on some days. The comparable results found between these studies, during follow-up, suggest fading may be a successful strategy for removing the intervention without compromising improvements in behaviour. This aligns with other research, which has found fading procedures successful at maintaining improvements in target behaviours (Liu et al., 2015; Petursdottir & Ragnarsdottir, 2019). Further research to investigate the effects of fading in a tootling procedure is recommended.

Teacher 1B in the current study continued to implement tootling, in Group 3, for 2 weeks before stopping, and was not using tootling during follow-up, three weeks later. This group had the biggest decline in on-task behaviour during the second follow-up session, which is more consistent with previous reports of a lack of maintenance when teachers chose not to continue with tootling implementation (Lum et al., 2017). Results from Group 1, which was exposed to tootling for longer than any other groups in the current study and those reported previously, contradict those of earlier studies (Lambert et al., 2015; Lum et al., 2017), as on-task behaviour remained high during follow-up, yet tootling in the classroom stopped immediately following the intervention phase. This result is interesting because maintenance of behaviour usually requires contact with naturally occurring reinforcement (Stokes & Baer, 1977). Therefore, with the removal of the tootling components, and no change to the rate of teacher praise statements, there is no apparent function of maintained rates of on-task behaviour, which aligns with results from Group 3, showing a decline in on-task behaviour once tootling was removed. However, it is possible that student behaviour changed in ways not related to behaviours measured in the current study because of the longer exposure to the intervention. Like findings which suggest tootling can decrease antisocial behaviour (Kirkpatrick et al., 2019), it is possible that tootling may have

increased prosocial behaviour in Group 1, therefore exposing students to natural reinforcers within the classroom, as students learned to be kinder to each other, maintaining rates of on-task behaviour. More research is needed to explore this theory. Regardless, the results from the current study indicate that longer exposure to tootling relates to more maintenance, and that a fading procedure such as the one implemented by Teacher 2 could be a promising way of removing the intervention without compromising results.

Students in Group 2 had the option of writing their own tootles or ticking a box on a pre-written tootle, because not all students were confident writers, and it was important that this skill deficit did not interfere with the procedure. The third data point for Group 2 in the tootling phase is the lowest in that phase and the only one to overlap with baseline data. This session coincided with the teacher implementing a change to the tootling procedure the previous day, which lasted for two days, whereby she instructed students to write more meaningful tootles to earn double-points. The third data point in Figure 2 equates to day 18 in Figure 1, which shows this day and the day before had the least number of tootles collected by this group. It is possible that the teacher's adaptation affected the number of tootles written, alongside on-task behaviour, during this time, as both variables increased immediately once the adaptation was removed. Dillon et al. (2019) reported that teachers in their study also questioned the content of the tootles and discovered students tootling on peers who were absent from class that day, and themselves. However, they also found improvements in behaviour, regardless of these discrepancies. Similar results have been found during implementation of other interventions designed to improve classroom behaviour, such as self-management procedures, whereby the accuracy of self-monitoring has not been critical for increasing on-task behaviour of some students (Graham-Day et al., 2010). Derieux (2019) investigated the effects of different written components of tootling and found that the content of the tootles had little effect on behaviour in their study. Together,

these studies, along with the changes in behaviour when the teacher adapted the intervention in this study, suggest that classroom behaviour improves as a result of tootling, regardless of the content of the tootles, indicating that other components, may be responsible for behaviour change. It is more likely that the consequences of behaviour, such as teacher praise and public display of progress are the fundamental reinforcing variables in the tootling procedure, which aligns with basic ABA principles and other research which has found functional relationships between increased teacher praise statements and improved classroom behaviour (Chalk & Bizo, 2004; O' Handley et al., 2020) as well as public posting and appropriately engaged behaviours (O' Handley et al., 2020).

Individual Student from Group 2. During Group 2's classroom observations, it became apparent that Nikau's behaviour dramatically improved once the tootling intervention was introduced. Therefore, Nikau's on-task behaviour was graphed separately to assess the effects of tootling on a student showing signs of at-risk behaviour. Visual analysis of Nikau's results show a sudden increase in on-task behaviour during the intervention phase, which continued throughout this phase and follow-up. Studies on tootling, which have previously examined the behaviour of target students have found similar results (Lipscomb et al., 2018; McHugh et al., 2016). The natural fading procedure implemented by Teacher 2 in this group, coupled with Nikau's maintained improvements in on-task behaviour, indicate that tootling could be useful as a secondary tier intervention for students at-risk, without the need for removing these students from the classroom to participate in small group interventions. Tootling could potentially improve behaviour of target students and, alongside improvements to classroom environments, reduce the possibility of social isolation (Skinner et al., 2000). However, data analysis of Nikau's behaviour was not part of the original study design and was only included in retrospect. Therefore, once observations were complete, no repeated measures were taken, meaning the methods were more like a case study than an

experimental design. For this reason, further exploring of the effects of tootling on the behaviour of students at-risk of emotional and behavioural disorder is recommended.

Question 2

Question 2 examined the impact of a tootling intervention on class-wide student disruptions of year 1 and year 5/6 primary-school classrooms. Visual analysis shows a decrease in student disruptions during the tootling phase, compared to baseline, for Group 2, which is consistent with previous investigations (Cihak et al., 2009; Dillon et al., 2019; Lambert et al., 2015; Lum et al., 2019; Lum et al., 2017; McHugh et al., 2016). A very large effect size, as demonstrated by a Tau-U score of .92 supports this result. Similar Tau-U scores were found by others using the same effect size calculation (Lum et al., 2019; Lum et al., 2017). Disruptions remained low for Group 2 during the follow-up phase, which is consistent with findings reported by Lambert et al. (2015), but different from results found by Lum et al. (2017). Like on-task behaviour, these differences are possibly because Teacher 2 continued to use tootling after the intervention phase, as did the teachers in the Lambert et al. study and unlike the teachers in the Lum et al. (2017) study.

For Groups 1 and 3, little change to student disruptions during the tootling phase, compared to baseline, was found during visual analysis, which was supported by small effect sizes. Therefore, it is unlikely there was a functional relationship between tootling and student disruptions for these groups, suggesting no causal effects can be established. These findings are contrary to results of previous studies (Cihak et al., 2009; Dillon et al., 2019; Lambert et al., 2015; Lum et al., 2019; Lum et al., 2017; McHugh et al., 2016). Regarding the differences found between groups within the current study, student disruptions during baseline for Groups 1 and 3 were generally much lower than for Group 2, with little change during tootling. Age of students may account for these differences. Students in Groups 1 and 3 were 10- to 11-years-old, whereas students in Group 2 were 5- 6-years-old. When students

are exposed to school- and class-wide punishment strategies, they often learn to avoid punishment by not "getting caught" by the teacher (Cashwell et al., 2001; Skinner et al., 2000). The low rates of on-task behaviour during baseline, compared to intervention, of the current study, alongside field notes indicating that students in Groups 1 and 3 often participated in passive-off-task behaviour suggest that the older students may have learned to quietly avoid engaging in academic materials, thus not drawing the teacher's attention. Conversely, younger students have not yet learned this behaviour, therefore, engage in more disruptions, which are more likely to elicit teacher corrections. Although this behaviour was not captured by the functional definition used during the current study, in one study which did measure passive-off-task behaviour, no change was found during intervention phases, compared to baseline and withdrawal phases (Lum et al., 2019). However, accurately capturing passive-off-task behaviour could be difficult, using a topographical definition, as used by Lum et al., because of difficulty defining the form of the behaviour. A functional definition which considers academic output might be a better option for measuring this behaviour, and future research which investigates this could provide interesting results.

Although previous researchers have found decreases in rates of disruptive behaviour across a variety of ages, it is difficult to compare those results to the findings in the current study because of the difference in operational definitions of disruptions. Previous literature, which has observed effects of an intervention on disruptive behaviour, has included topographical definitions (Cihak et al., 2009; Dillon et al., 2019; Lambert et al., 2015; Lum et al., 2019; Lum et al., 2017; McHugh et al., 2016; Rubow et al., 2018), whereas the definition in the current study was a functional one, based on the rate of teacher corrections. This functional definition was chosen for two reasons. Firstly, measuring whole-class behaviour, using a topographical definition would have been difficult in practice, whereas a functional definition made it feasible. Secondly, it considered the effect behaviour had on the

environment, instead of simply describing its form which, ultimately, may or may not have disrupted the flow of the classroom, as perceived by the teacher. For example, in their study on the effects of a self-monitoring procedure on on-task behaviour, Graham-Day et al. (2010) described their topographical definition of behaviour as a limitation of their study after observing one participant stay on-task, but out of her seat for the entire lesson and recording her behaviour as not on-task because of the restrictions caused by the operational definition. Similarly, previous researchers, using topographical definitions of disruptive behaviour, have identified the need to code appropriate and disruptive behaviour simultaneously, on occasions when students are engaging in both, such as working on-task, whilst talking to a peer, for example (Dillon et al., 2019). However, if students are on-task, the teacher may choose to allow work-related peer discussions, instead of requesting help from the teacher, and believe it to be more helpful than disruptive. Coding behaviour simultaneously may lead to an over-representation of behaviour. Classrooms are fluid environments, and rules can change during a lesson, depending on how the teacher perceives each situation, as it evolves. Using a functional definition of student disruptions, measured using teacher corrections, instead of a topographical definition of behaviour which may, or may not, disrupt the classroom can provide researchers with a more accurate measure with which to draw conclusions.

Question 3

The third research question addressed in this study considered the effect of the tootling intervention on teacher praise statements. Results from the teachers across all three classrooms indicate that praise statements did not increase because of tootling. Although the current study was the first of its kind to measure the effects of a tootling intervention on rates of teacher praise, others have examined the effects of The Good Behaviour Game on student and teacher behaviour (Elswick & Casey, 2011; Lannie & McCurdy, 2007; Rubow etal., 2018). Results from the current study were like those found by Lannie and

McCurdy (2007), who reported teacher rates of praise remained at near zero across all phases of a withdrawal design. Lannie and McCurdy argued that student disruptive behaviour is more likely to elicit a response from the teacher than passive behaviour, such as on-task behaviour. The results from the current study support this claim, as on-task behaviour increased during the tootling phase, alongside a decrease in student disruptions (measured by teacher corrective statements), but praise statements remained low, regardless of the increased opportunities for the teachers to respond positively to behaviour. A behavioural explanation for this lack of praise relates to reinforcement theory. Any behaviour which is followed by reinforcement is more likely to occur again in future (Skinner, 1968). When a teacher responds to student disruption, that disruptive behaviour will likely cease, temporarily, thus providing the teacher with negative reinforcement. However, when a teacher responds to appropriate behaviour, there is no change in the environment and therefore nothing to reinforce the teacher's behaviour. With nothing to prompt the teacher to respond and no history of reinforcement for doing so, the behaviour is unlikely to re-occur (Skinner, 1968). This idea is supported by previous research, which found that only when reinforcement was contingent on increased levels of teacher praise, did teacher praise increase (Eaves, 2020), alongside findings that behaviour-specific praise increased during observation feedback (reinforcement) phases but was not maintained during withdrawal phases (Sutherland et al., 2000).

The current praise data indicate that tootling had a small counter-therapeutic effect on praise statements for Teachers 1A and 1B. Alongside this, Teacher 2 had the highest levels of praise statements, alongside high levels of student disruptions. This is interesting because, according to the basic principles of reinforcement, praise should reinforce appropriate behaviour, therefore reducing inappropriate behaviour (Ferster & Skinner, 1957). However, these results could indicate that disruptive behaviour not only elicits teacher responding in the

form of corrective statements, as discussed above, but responding in the form of praise, too. For example, high levels of student disruptions may prompt teacher behaviour and be a discriminative stimulus for praise behaviour, because praise results in decreased rates of disruptions. However, when rates of student disruptions are lower, there is no prompt and no discriminative stimulus indicating that reinforcement is available for praise behaviour. Unfortunately, providing students with more praise when they are being disruptive, than when they are on-task, in effect, reinforces disruptive behaviour, and could contribute to the explanation of why behaviour change in previous studies was not maintained once behaviour interventions were removed from the classroom (Cihak et al., 2009; Dillon et al., 2019; McHugh et al., 2016).

Although results from the current study coincide with results found by Lannie and McCurdy (2007), they are dissimilar to findings reported by others. Rubow et al. (2018) and Elswick and Casey (2011) found praise statements increased during intervention phases, compared to baseline and withdrawal phases. However, Rubow et al. included specific praise statements related to academic work in their operational definitions of praise, which were excluded from the definition in the current study. Therefore, it is unclear whether the Rubow et al. findings were related directly to classroom behaviour or academic responding. Teachers have been found to naturally praise academic student responses more than behaviour (Harrop & Swinson, 2000), meaning it is important to distinguish between the two because, whilst praising an academic response will likely increase academic responding in future, it is unlikely to affect other classroom behaviour. Although there is no clear explanation as to why Elswick and Casey (2011) found differing results to those reported in the current study, it should be noted that their study used an AB design, with both teacher participants moving into the intervention phase at the same time and no withdrawal or re-implementation phases.

Without this level of control, it is difficult to draw conclusions regarding any functional relationships between the dependent and independent variables.

Although results from the current study indicate observed rates of teacher praise did not increase during the intervention phase, teacher praise, contingent on prosocial behaviour, is a component of the tootling procedure. At the end of each day (or mathematics lesson) teachers provided a selection of students with praise, contingent on prosocial behaviour, as reported by peers. This additional daily praise likely reinforced prosocial behaviour, thus contributing to appropriate classroom behaviour. Therefore, in classrooms with low rates of teacher praise, tootling could be utilised as a strategy to prompt this behaviour.

Question 4

The final research question addressed the social validity of the tootling intervention. Based on the scores of the modified BIRS, teachers across all groups found tootling to be an acceptable intervention for improving classroom behaviour. These results are consistent with previous studies which used the modified BIRS to assess social validity of a tootling intervention (Dillon et al., 2019; Lipscomb et al., 2018; Lum et al., 2019; Lum et al., 2017). Also consistent with results found by the above researchers, is the consistently lower scores found in relation to the effectiveness factor, than the acceptability and time to effect factor (Dillon et al., 2019; Lum et al., 2019; Lum et al., 2017). This is interesting considering the substantial improvements in on-task behaviour in the current study, and academically engaged behaviour in the previous studies. These findings suggest that teachers are unaware of improvements in classroom behaviour, which is supported by the continued low rates of praise even after student behaviour has improved. Thus, teachers are neither praising appropriate behaviour, nor aware of it, which supports claims that teachers could benefit from factual, objective feedback in real time (Sutherland et al., 2000). It is also likely that an

intervention such as tootling could be beneficial in classrooms, as it includes a praise component, removing the need to rely on teachers to naturally praise behaviour.

Alongside the teachers' ratings of acceptability, the students also voluntarily completed a modified CIRP, which resulted in all, but two, students rating the tootling intervention as acceptable. Again, these findings are consistent with previous research (Lipscomb et al., 2018; Lum et al., 2019). It is worth noting that in Group 2, the teacher suggested that questions requiring reverse scoring were confusing for the students, who did not understand how to respond appropriately. This could explain why so many students (53%) in this group did not complete all the questions. This issue has not been reported in the tootling literature before. However, Lum et al. (2019) did highlight a modification to their CIRP which involved reverse scoring. This modification, therefore, likely improved response rate in that study, compared to the current study.

Measures of Integrity

Procedural integrity was 100% during teacher and student training, across groups. These results are consistent with previous research and indicate that tootling is a simple enough procedure, which teachers can learn and deliver to students, without requiring multiple training sessions. Teacher reported treatment integrity remained at 94%, or above, throughout the tootling phase, while mean observer scores were lower for Teacher 1B (75%), and similar for Teacher 1A (96%) and Teacher 2 (100%). Some researchers have reported lower levels of treatment integrity for teacher participants previously (Lambert et al., 2015; Lum et al., 2017; McHugh et al., 2016). However, as noted by others, levels of improvement regarding classroom behaviour indicate that some intervention steps can be missed with little detriment (Lambert et al., 2015; McHugh et al., 2016).

Daily tootle count was also recorded as a measure of treatment integrity. Students did consistently report peers' prosocial behaviour during the tootling phase, which is consistent

with previous studies (Cashwell et al., 2001; Derieux, 2019; Skinner et al., 2000) and provides further evidence of a functional relation between daily tootles and classroom behaviour.

Limitations and Future Research

While the current study produced some promising results regarding tootling as successful classroom intervention, there are some limitations which should be considered. Like other research on tootling, data was collected from three groups of students, across two age-groups, in one school, meaning generalisability of results to students in other schools and settings is limited. Although the current study adds to previous investigations on tootling, which have examined the effects of tootling on various behaviours of students across age-groups, with and without intellectual difficulties, in different settings (Dillon et al., 2019; Kirkpatrick et al., 2019; Lambert et al., 2015; Lipscomb et al., 2018; Lum et al., 2019; McHugh et al., 2016), more replications are needed to strengthen the validity of the research. Particularly, more research is recommended on the effects of tootling on teacher behaviour, as the current study was the first of this kind. Importantly, since the participating school in the current study was a PB4LSW school, behaviour of teachers from schools that are not PB4LSW schools should be examined in relation to tootling, as it is unclear what effect this framework had on behaviour before commencement of the research. Although, considering the rates of teacher praise during baseline, it is suspected that little difference would be found prior to intervention implementation.

Although classroom behaviour improved because of the tootling intervention, some of the lower treatment integrity scores in the current study, as seen in previous studies, shouldbe addressed. In line with previous research, the data around treatment integrity indicates thatthe steps missed by the teacher were those requiring teacher action, such as counting and reading tootles, and updating the tootle chart. Lum et al. (2019) addressed this issue by

implementing an independent group contingency in place of the interdependent group contingency used in the current study, which removed some of the steps to be completed by the teacher. However, previous researchers have found interdependent group contingencies more successful at improving behaviour than independent group contingencies (Gresham & Gresham, 1982), indicating independent group contingencies may not be an ideal option. Research shows peers can be trained in roles usually reserved for teachers (Harjusola-Webb et al., 2012; Kaya et al., 2015), as the current study demonstrates. Therefore, future researchers could incorporate procedures which utilise students more in tootling tasks. For example, students could be selected to perform different tasks, usually performed by the teacher, each day. Allocating responsibilities such as these to students, instead of teachers, may free-up time for teachers to focus on academic instruction, potentially increasing social validity, alongside treatment integrity.

Another limitation of the current study relates to the teacher participants and student groups. Two of the groups were “home” classes, together in the classroom for much ofthe day, whilst Group 3 was a mathematics class, together for one lesson. Likewise, there were students in Group 1 that were also in Group 3. Overall, this meant that the students in each group were not exposed to tootling equally. Regardless, results from the multiple- baseline design, which provided repeated measures to establish reliability of effect, suggest that on-task behaviour changed in line with the implementation of the tootling intervention within each group, suggesting a causal relationship between tootling and behaviour change. Similarly, although Teacher 1 taught Groups 1 and 3, there is no evidence to suggest a change in behaviour whilst teaching Group 3 because of the intervention being introduced to Group 1. Regardless, future researchers faced with similar situations, should consider waysto increase internal validity, such as including a reversal element within their experimental design.

Results show that the current study did not improve rates of teacher behaviour. Eaves (2020) argued that reinforcement, contingent on increased rates of teacher praise is necessary to increase praise, even if teachers understand the benefits of praising appropriate behaviour. This idea is supported by basic reinforcement principles (Ferster & Skinner, 1957; Skinner, 1968). Although tootling incorporates teacher praise into its procedures, an increase in natural teacher praise statements can generate greater improvements to classroom environments. Future researchers could investigate ways to incorporate the reinforcement of teacher praise into the tootling procedure. For example, teachers could be included in the tootling intervention by training students to observe and report teacher behaviour alongside peer prosocial behaviour. Sutherland et al. (2000) argued that continuous feedback may be needed to maintain high levels of behaviour-specific praise. Students reporting positive teacher behaviour could offer a solution to this need for continuous feedback, thus resulting in the reinforcement of teacher praise statements as supported by Eaves. Introducing an element which focusses on improving teacher behaviour, to an intervention which already improves student behaviour, could contribute to a more positive classroom environment whilst saving valuable school resources.

In all the tootling literature, including the results from the current study, no one has reported the effects of tootling on prosocial behaviour. Following on from the original studies, which reported that an interdependent group contingency, alongside public display of progress increased reports on peers' prosocial behaviour (Cashwell et al., 2001; Skinner et al., 2000), the majority of research has investigated the effects of tootling on behaviour related to academic engagement and disruptions (Cihak et al., 2009; Dillon et al., 2019; Lambert et al., 2015; Lipscomb et al., 2018; Lum et al., 2019; Lum et al., 2017; McHugh et al., 2016), with the exception of one study which examined effects on antisocial behaviour (Kirkpatrick et al., 2019). It is likely that teacher praise for prosocial behaviour reported by

peers, because of tootles being read aloud and praised, reinforces prosocial, as well as on-task, behaviour. This could explain maintenance of behaviour in Group 1, as discussed previously. However, more research is needed to investigate these claims. This is especially important since programmes designed to teach social skills away from the classroom have been found to lack generalisation to the natural environment (Cashwell et al., 2001; Stokes & Baer, 1977).

Implications for Practice and Conclusion

Schools need empirically supported interventions which are simple to use and time efficient, to promote positive classroom behaviour. Tootling, alongside an interdependent group contingency and public display of progress shows promise at meeting these criteria. Tootling can be used as a proactive, primary tier, classroom support in junior and senior primary-school classrooms, as well as with students with specific academic skill deficits, to help teach and reinforce appropriate classroom behaviour. Continued use of the tootling procedure should result in high rates of on-task behaviour, which are maintained over time. Tootling may also function as a secondary tier intervention, benefitting students at risk, as seen from Nikau's results, alongside maintained results following Teacher 2's natural fading procedure. Although tootling did not result in an increase in naturally occurring teacher praise statements, the praise component could contribute to this area, without the requirement of extra time and resources needed for further teacher training. With the shift towards evidence-based, positive behaviour interventions in schools, more research on interventions such as tootling and the effects on student and teacher behaviour is necessary. Tools such as this are essential for schools to provide supportive classroom environments which benefit both the students and teachers within them.

References

Allday, R. A., Hinkson-Lee, K., Hudson, T., Neilsen-Gatti, S., Kleinke, A., & Russel, C. S. (2012). Training general educators to increase behavior-specific praise: Effects on students with ebd. *Behavioral Disorders, 37*(2), 87-98. https://doi.org/10.1177/019874291203700203

Aloe, A. M., Shisler, S. M., Norris, B. D., Nickerson, A. B., & Rinker, T. W. (2014). A multivariate meta-analysis of student misbehavior and teacher burnout. *Educational Research Review, 12*(C), 30-44. https://doi.org/10.1016/j.edurev.2014.05.003

Alvero, A. M., Struss, K., & Rappaport, E. (2008). Measuring safety performance: A comparison of whole, partial, and momentary time-sampling recording methods. *Journal of Organizational Behavior Management, 27*(4), 1-28. https://doi.org/10.1300/J075v27n04_01

Antoniou, P., & Kyriakides, L. (2013). A Dynamic Integrated Approach to teacher professional development: Impact and sustainability of the effects on improving teacher behaviour and student outcomes. *Teaching and Teacher Education, 29*(1), 1-12. https://doi.org/10.1016/j.tate.2012.08.001

Baer, D. M. (1999). *How to plan for generalization* (2nd ed.). Pro-Ed.

Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 1*(1), 91-97. https://doi.org/10.1901/jaba.1968.1-91

Bambara, L. M., Cole, C. L., Kunsch, C., Tsai, S.-C., & Ayad, E. (2016). A peer-mediated intervention to improve the conversational skills of high school students with Autism Spectrum Disorder. *Research in Autism Spectrum Disorders, 27*, 29-43. https://doi.org/10.1016/j.rasd.2016.03.003

Barrish, H. H., Saunders, M., & Wolf, M. M. (1969). Good behavior game: Effects of individual contingencies for group consequences on disruptive behavior in a classroom. *Journal of Applied Behavior Analysis, 2*(2), 119-124. https://doi.org/10.1901/jaba.1969.2-119

Barth, J. M., Dunlap, S. T., Dane, H., Lochman, J. E., & Wells, K. C. (2004). Classroom environment influences on aggression, peer relations, and academic focus. *Journal of School Psychology, 42*(2), 115-133. https://doi.org/10.1016/j.jsp.2003.11.004

Boniecki, K. A., & Moore, S. (2003). Breaking the silence: Using a token economy to reinforce classroom participation. *Teaching of Psychology, 30*(3), 224-227. https://doi.org/10.1207/S15328023TOP3003_05

Bonik, K. (2018). Teacher's behaviour towards students' motivation practice. *SHS Web of Conferences, 42*, 1-6. https://doi.org/10.1051/shsconf/20184200078

Bost, L. W., & Riccomini, P. J. (2006). Effective instruction: An inconspicuous strategy for dropout prevention. *Remedial and Special Education, 27*(5), 301-311. https://doi.org/10.1177/07419325060270050501

Bowman-Perrott, L., Burke, M. D., Zaini, S., Zhang, N., & Vannest, K. (2016). Promoting positive behavior using the good behavior game: A meta-analysis of single-case research. *Journal of Positive Behavior Interventions, 18*(3), 180-190. https://doi.org/10.1177/1098300715592355

Brogan, K. M., Falligant, J. M., & Rapp, J. T. (2017). Interdependent group contingencies decrease adolescents' disruptive behaviors during group therapy: A practitioner's demonstration. *Behavior Modification, 41*(3), 405-421. https://doi.org/10.1177/0145445517693812

Camacho, R., Anderson, A., Moore, D., & Furlonger, B. (2014). Conducting a function-based intervention in a school setting to reduce inappropriate behaviour of a child with autism. *Behaviour Change, 31*(1), 65-77. https://doi.org/10.1017/bec.2013.33

Cashwell, T., Skinner, C., & Smith, E. (2001). Increasing second-grade students' reports of peers' prosocial behaviors via direct instruction, group reinforcement, and progress feedback: A replication and extension. *Education & Treatment of Children, 24*(2), 161.

Chaffee, R., Briesch, A., Johnson, A., & Volpe, R. (2017). A meta-analysis of class-wide interventions for supporting student behavior. *School Psychology Review, 46*(2), 149-164. https://doi.org/10.17105/SPR-2017-0015.V46-2

Chalk, K., & Bizo, L. A. (2004). Specific praise improves on-task behaviour and numeracy enjoyment: A study of year four pupils engaged in the numeracy hour. *Educational Psychology in Practice, 20*(4), 335-351. https://doi.org/10.1080/0266736042000314277

Cihak, D. F., Kirk, E. R., & Boon, R. T. (2009). Effects of classwide positive peer "tootling" to reduce the disruptive classroom behaviors of elementary students with and without disabilities. *Journal of Behavioral Education, 18*(4), 267-278. https://doi.org/10.1007/s10864-009-9091-8

Coogan, B. A., Kehle, T. J., Bray, M. A., & Chafouleas, S. M. (2007). Group contingencies, randomization of reinforcers, and criteria for reinforcement, self-monitoring, and peer feedback on reducing inappropriate classroom behavior. *School Psychology Quarterly, 22*(4), 540-556. https://doi.org/10.1037/1045-3830.22.4.540

Cooper, J. O., Heron, T. E., & Heward, W. L. (2014). *Applied behavior analysis* (Second ed.). Pearson Education.

Cooper, P., & Jacobs, B. (2011). Pupils making a difference: Enhancing the power of the student peer group to promote positive social, emotional and behavioural outcomes. *Emotional and Behavioural Difficulties*, *16*(1), 5-13. https://doi.org/10.1080/13632752.2011.545594

Dart, E. H., Collins, T. A., Klingbeil, D. A., & McKinley, L. E. (2014). Peer management interventions: A meta-analytic review of single-case research. *School Psychology Review*, *43*(4), 367-384. https://doi.org/10.17105/SPR-14-0009.1

Derieux, J. (2019). *Is the written component of tootling effective? A comparison of the group contingency to a comparison writing procedure* [Master's thesis, University of Southern Mississippi]. Aquila. https://aquila.usm.edu/masters_theses/712

Dillon, M. B. M., Radley, K., Tingstrom, D., Dart, E., & Barry, C. (2019). The effects of tootling via classdojo on student behavior in elementary classrooms. *School Psychology Review*, *48*(1), 18-30. https://doi.org/10.17105/SPR-2017-0090.V48-1

Eaves, A. (2020). *Using two reinforcement contingencies to alter teachers' rates of praise* [Doctoral dissertation, The University of Southern Mississippi]. Aquila. https://aquila.usm.edu/dissertations/1798

Eaves, A. E., Radley, K. C., Dufrene, B. A., Olmi, D. J., & Bernard, L. E. (2020). A comparison of two group contingencies on teachers' use of behavior-specific praise. *Journal of Behavioral Education*. https://doi.org/10.1007/s10864-020-09366-6

Elder, K., & Prochnow, J. (2016). PB4L school-wide: What will support the sustainability of the initiative? *New Zealand Journal of Educational Studies*, *51*(1), 83-97. https://doi.org/10.1007/s40841-016-0036-1

Elliott, S. N. (1986). Children's ratings of the acceptability of classroom interventions for misbehavior: Findings and methodological considerations. *Journal of School Psychology*, *24*(1), 23-35. https://doi.org/10.1016/0022-4405(86)90039-7

Elliott, S. N., & Treuting, M. V. B. (1991). The behavior intervention rating scale: Development and validation of a pretreatment acceptability and effectiveness measure. *Journal of School Psychology, 29*(1), 43-51. https://doi.org/10.1016/0022-4405(91)90014-I

Elswick, S., & Casey, L. B. (2011). The good behavior game is no longer just an effective intervention for students: An examination of the reciprocal effects on teacher behaviors. *Beyond Behavior, 21*(1), 36-46.

Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement*. Prentice-Hall.

Fettig, A., & Barton, E. E. (2014). Parent implementation of function-based intervention to reduce children's challenging behavior: A literature review. *Topics in Early Childhood Special Education, 34*(1), 49-61. https://doi.org/10.1177/0271121413513037

Filcheck, H. A., McNeil, C. B., Greco, L. A., & Bernard, R. S. (2004). Using a whole-class token economy and coaching of teacher skills in a preschool classroom to manage disruptive behavior. *Psychology in the Schools, 41*(3), 351-361. https://doi.org/10.1002/pits.10168

Gage, N. A., Sugai, G., Lewis, T. J., & Brzozowy, S. (2015). Academic achievement and school-wide positive behavior supports. *Journal of Disability Policy Studies, 25*, 199-209. https://doi.org/10.1177/1044207313505647

Goss, A., Lynette. (2019). *The effects of tootling without external rewards on whole-class lessons on a lower-elementary classroom* [Masters thesis, The Saint Catherine University]. Sophia. https://sophia.stkate.edu/maed/339

Graham-Day, K. J., Gardner, R., & Y, H. (2010). Increasing on-task behaviors of high school students with attention deficit hyperactivity disorder: Is it enough? *Education & Treatment of Children, 33*(2), 205-221. https://doi.org/10.1353/etc.0.0096

Grauvogel-Macaleese, A. N., & Wallace, M. D. (2010). Use of peer-mediated intervention in children with attention deficit hyperactivity disorder. *Journal of Applied Behavior Analysis, 43*(3), 547-551. https://doi.org/10.1901/jaba.2010.43-547

Gresham, F. M. (2005). Treatment integrity and therapeutic change: Commentary on perepletchikova and kazdin. *Clinical Psychology: Science and Practice, 12*(4), 391-394. https://doi.org/10.1093/clipsy.bpi048

Gresham, F. M., Cook, C. R., Crews, S. D., & Kern, L. (2004). Social skills training for children and youth with emotional and behavioral disorders: Validity considerations and future directions. *Behavioral Disorders, 30*(1), 32-46. https://doi.org/10.1177/019874290403000101

Gresham, F. M., & Gresham, G. N. (1982). Interdependent, dependent, and independent group contingencies for controlling disruptive behavior. *Journal of Special Education, 16*(1), 101-110. https://doi.org/10.1177/002246698201600110

Grieger, T., Kauffman, J. M., & Grieger, R. M. (1976). Effects of peer reporting on cooperative play and aggression of kindergarten children. *Journal of School Psychology, 14*(4), 307-313. https://doi.org/10.1016/0022-4405(76)90027-3

Groves, E. A., & Austin, J. L. (2017). An evaluation of interdependent and independent group contingencies during the good behavior game. *Journal of Applied Behavior Analysis, 50*(3), 552-556. https://doi.org/10.1002/jaba.393

Gunter, P., Venn, M., Patrick, J., Miller, K., & Kelly, L. (2003). Efficacy of using momentary time samples to determine on-task behavior of students with emotional/behavioral disorders. *Education & Treatment of Children, 26*(4), 400-412.

Harbour, K. E., Evanovich, L. L., Sweigart, C. A., & Hughes, L. E. (2015). A brief review of effective teaching practices that maximize student engagement. *Preventing School*

Failure: Alternative Education for Children and Youth, 59(1), 5-13. https://doi.org/10.1080/1045988X.2014.919136

Harjusola-Webb, S., Parke Hubbell, S., & Bedesem, P. (2012). Increasing prosocial behaviors of young children with disabilities in inclusive classrooms using a combination of peer-mediated intervention and social narratives. *Beyond Behavior, 21*(2), 29-36.

Harrop, A., & Swinson, J. (2000). Natural rates of approval and disapproval in British infant, junior and secondary classrooms. *British Journal of Educational Psychology, 70*(4), 473-483. https://doi.org/10.1348/000709900158236

Hendrickx, M. M. H. G., Mainhard, T., Oudman, S., Boor-Klip, H. J., & Brekelmans, M. (2017). Teacher behavior and peer liking and disliking: The teacher as a social referent for peer status. *Journal of Educational Psychology, 109*(4), 546-558. https://doi.org/10.1037/edu0000157

Henington, C., & Skinner, C. H. (1998). Peer monitoring. In K. Topping & E. Stewart (Eds.), *Peer Assisted Learning* (pp. 237-253). Routledge.

Jalilifar, A. (2010). The effect of cooperative learning techniques on college students' reading comprehension. *System, 38*(1), 96-108. https://doi.org/10.1016/j.system.2009.12.009

Johansen, A., Little, S. G., & Akin-Little, A. (2011). An examination of New Zealand teachers' attributions and perceptions of behaviour, classroom management, and the level of formal teacher training received in behaviour management. *Kairaranga, 12*(2), 3-12.

Johnson, D. W., Johnson, R. T., & Roseth, C. (2010). Cooperative learning in middle schools interrelationship of relationships and achievement. *Middle Grades Research Journal, 5*(1), 1-18.

Joslyn, P., Vollmer, T., & Kronfli, F. (2019). Interdependent group contingencies reduce disruption in alternative high school classrooms. *Journal of Behavioral Education, 28*(4), 423-434. https://doi.org/10.1007/s10864-019-09321-0

Kaya, C., Blake, J., & Chan, F. (2015). Peer-mediated interventions with elementary and secondary school students with emotional and behavioural disorders: A literature review. *Journal of Research in Special Educational Needs, 15*(2), 120-129. https://doi.org/10.1111/1471-3802.12029

Kennedy, C. H. (2005). *Single-case designs for educational research*. Pearson.

Kirkpatrick, B. A., Wright, S., Daniels, S., Taylor, K. L., McCurdy, M., & Skinner, C. H. (2019). Tootling in an after-school setting: Decreasing antisocial interactions in at-risk students. *Journal of Positive Behavior Interventions, 21*(4), 228-237. https://doi.org/10.1177/1098300719851226

Kohler, F. W., & Strain, P. S. (1990). Peer-assisted interventions: Early promises, notable achievements, and future aspirations. *Clinical Psychology Review, 10*(4), 441-452. https://doi.org/10.1016/0272-7358(90)90047-E

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). *Single-case designs technical documentation*. What Works Clearinghouse. http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf

Kyndt, E., Raes, E., Lismont, B., Timmers, F., Cascallar, E., & Dochy, F. (2013). A meta-analysis of the effects of face-to-face cooperative learning. Do recent studies falsify or verify earlier findings? *Educational Research Review, 10*, 133-149. https://doi.org/10.1016/j.edurev.2013.02.002

Kyriakides, L., Creemers, B. P. M., & Antoniou, P. (2009). Teacher behaviour and student outcomes: Suggestions for research on teacher training and professional development.

Teaching and Teacher Education, 25(1), 12-23. https://doi.org/10.1016/j.tate.2008.06.001

Lambert, A. M., Tingstrom, D. H., Sterling, H. E., Dufrene, B. A., & Lynne, S. (2015). Effects of tootling on classwide disruptive and appropriate behavior of upper-elementary students. *Behavior Modification, 39*(3), 413-430. https://doi.org/10.1177/0145445514566506

Lane, J. D., & Gast, D. L. (2013). Visual analysis in single case experimental design studies: Brief review and guidelines. *Neuropsychol Rehabil, 24*(3-4), 445-463. https://doi.org/10.1080/09602011.2013.815636

Lane, K. L., Menzies, H. M., Ennis, R. P., & Bezdek, J. (2013). School-wide systems to promote positive behaviors and facilitate instruction. *Journal of Curriculum and Instruction, 7*(1), 6-31.

Lannie, A., & McCurdy, B. (2007). Preventing disruptive behavior in the urban classroom: Effects of the good behavior game on student and teacher behavior. *Education & Treatment of Children, 30*(1), 85-98. https://doi.org/10.1353/etc.2007.0002

Lassen, S. R., Steele, M. M., & Sailor, W. (2006). The relationship of school-wide Positive Behavior Support to academic achievement in an urban middle school. *Psychology in the Schools, 43,* 701-712. https://doi.org/10.1002/pits.20177

Lipscomb, A. H., Anderson, M., & Gadke, D. L. (2018). Comparing the effects of ClassDojo with and without Tootling intervention in a postsecondary special education classroom setting. *Psychology in the Schools, 55*(10), 1287-1301. https://doi.org/10.1002/pits.22185

Liu, Y., Moore, D., & Anderson, A. (2015). Improving social skills in a child with autism spectrum disorder through self-management training. *Behaviour Change, 32*(4), 273-284. https://doi.org/10.1017/bec.2015.14

Lum, J. D. K. (2017). *Tootling with a randomized independent group contingency in a high school setting* [Masters thesis, The University of Southern Mississippi]. Aquila. https://aquila.usm.edu/dissertations/909

Lum, J. D. K., Radley, K. C., Tingstrom, D. H., Dufrene, B. A., Olmi, D. J., & Wright, S. J. (2019). Tootling with a randomized independent group contingency to improve high school classwide behavior. *Journal of Positive Behavior Interventions, 21*, 93-105. https://doi.org/10.1177/1098300718792663

Lum, J. D. K., Tingstrom, D. H., Dufrene, B. A., Radley, K. C., & Lynne, S. (2017). Effects of tootling on classwide disruptive and academically engaged behavior of general-education high school students. *Psychology in the Schools, 54*(4), 370-384. https://doi.org/10.1002/pits.22002

Lynch, D., & Keenan, M. (2018). The good behaviour game: Maintenance effects. *International Journal of Educational Research, 87*, 91-99. https://doi.org/10.1016/j.ijer.2016.05.005

Maggin, D. M., Chafouleas, S. M., Goddard, K. M., & Johnson, A. H. (2011). A systematic evaluation of token economies as a classroom management tool for students with challenging behavior. *Journal of School Psychology, 49*(5), 529-554. https://doi.org/10.1016/j.jsp.2011.05.001

Martin, J. E., & Rusch, F. R. (1987). Use of the partial-sequential withdrawal design to assess maintenance of mentally retarded adults' acquired meal preparation skills. *Research in Developmental Disabilities, 8*(3), 389-399. https://doi.org/10.1016/0891-4222(87)90021-7

McHugh, M. B., Tingstrom, D. H., Radley, K. C., Barry, C. T., & Walker, K. M. (2016). Effects of tootling on classwide and individual disruptive and academically engaged

behavior of lower-elementary students. *Behavioral Interventions, 31*(4), 332-354. https://doi.org/10.1002/bin.1447

Meany-Daboul, M. G., Roscoe, E. M., Bourret, J. C., & Ahearn, W. H. (2007). A comparison of momentary time sampling and partial-interval recording for evaluating functional relations. *Journal of Applied Behavior Analysis, 40*(3), 501-514. https://doi.org/10.1901/jaba.2007.40-501

Ministry of Education. (2019). *PB4L- School-Wide*. http://pb4l.tki.org.nz/PB4L-School-Wide

Ministry of Education. (2020). *School deciles*. https://www.education.govt.nz/school/funding-and-financials/resourcing/operational-funding/school-decile-ratings/

Mitchell, R. R., Tingstrom, D. H., Dufrene, B. A., Ford, W. B., & Sterling, H. E. (2015). The effects of the good behavior game with general-education high school students. *School Psychology Review, 44*(2), 191. https://doi.org/10.17105/spr-14-0063.1

Moore, D. W., Prebble, S., Robertson, J., Waetford, R., & Anderson, A. (2001). Self-recording with goal setting: A self-management programme for the classroom. *Educational Psychology, 21*(3), 255-265. https://doi.org/10.1080/01443410123466

Moroz, K., & Jones, K. (2002). The effects of positive peer reporting on children's social involvement. *School Psychology Review, 31*(2), 235-245.

Morrison, J., & Jones, K. (2007). The effects of positive peer reporting as a class-wide positive behavior support. *Journal of Behavioral Education, 16*(2), 111-124. https://doi.org/10.1007/s10864-006-9005-y

Mottram, A. M., Bray, M. A., Kehle, T. J., Broudy, M., & Jenson, W. R. (2002). A classroom-based intervention to reduce disruptive behaviors. *Journal of Applied School Psychology, 19*(1), 65-74. https://doi.org/10.1300/J008v19n01_05

Murphy, J., & Zlomke, K. (2014). Positive peer reporting in the classroom: A review of intervention procedures. *Behavior Analysis in Practice, 7*(2), 126-137.

Naeghel, J., Valcke, M., Meyer, I., Warlop, N., Braak, J., & Keer, H. (2014). The role of teacher behavior in adolescents' intrinsic reading motivation. *An Interdisciplinary Journal, 27*(9), 1547-1565. https://doi.org/10.1007/s11145-014-9506-3

Niesyn, M. E. (2009). Strategies for success: Evidence-based instructional practices for students with emotional and behavioral disorders. *Preventing School Failure: Alternative Education for Children and Youth, 53*(4), 227-234. https://doi.org/10.3200/PSFL.53.4.227-234

O' Handley, R. D., Olmi, D. J., Dufrene, B. A., Tingstrom, D. H., & Whipple, H. (2020). The effects of behavior-specific praise and public posting in secondary classrooms. *Psychology in the Schools, 57*(7), 1097-1115. https://doi.org/10.1002/pits.22375

Panayiotou, A., Kyriakides, L., Creemers, B., McMahon, L., Vanlaar, G., Pfeifer, M., Rekalidou, G., & Bren, M. (2014). Teacher behavior and student outcomes: Results of a European study. *Educational Assessment, Evaluation and Accountability, 26*(1), 73-93. https://doi.org/10.1007/s11092-013-9182-x

Parker, R. I., & Vannest, K. (2009). An improved effect size for single-case research: Nonoverlap of all pairs. *Behavior Therapy, 40*(4), 357-367. https://doi.org/10.1016/j.beth.2008.10.006

Parker, R. I., Vannest, K. J., Davis, J. L., & Sauber, S. B. (2011). Combining nonoverlap and trend for single-case research: Tau-U. *Behavior Therapy, 42*(2), 284-299. https://doi.org/10.1016/j.beth.2010.08.006.

Parsonage, M., Khan, L., & Saunders, A. (2014). *Building a better future: The lifetime costs of childhood behavioural problems and the benefits of early intervention*. Centre for Mental Health.

Parsonson, B. S. (2012). Evidence-based classroom behaviour management strategies. *Kairaranga, 13*(1), 16-23.

Petursdottir, A. L., & Ragnarsdottir, G. B. (2019). Decreasing student behavior problems and fostering academic engagement through function-based support and fading of token reinforcement. *Behavioral Interventions, 34*(3), 323-337. https://doi.org/10.1002/bin.1670

Plavnick, J. B., & Ferreri, S. J. (2013). Single-case experimental designs in educational research: A methodology for causal analyses in teaching and learning. *Educational Psychology Review, 25*(4), 1-21. https://doi.org/10.1007/s10648-013-9230-6

Reinke, W. M., Herman, K. C., & Stormont, M. (2013). Classroom-level positive behavior supports in schools implementing sw-pbis: Identifying areas for enhancement. *Journal of Positive Behavior Interventions, 15*(1), 39-50. https://doi.org/10.1177/1098300712459079

Reinke, W. M., Stormont, M., Herman, K. C., Wang, Z., Newcomer, L., & King, K. (2014). Use of coaching and behavior support planning for students with disruptive behavior within a universal classroom management program. *Journal of Emotional and Behavioral Disorders, 22*(2), 74-82. https://doi.org/10.1177/1063426613519820

Riley, J., McKevitt, B., Shriver, M., & Allen, K. (2011). Increasing on-task behavior using teacher attention delivered on a fixed-time schedule. *Journal of Behavioral Education, 20*(3), 149-162. https://doi.org/10.1007/s10864-011-9132-y

Ritzema, E. S., Deunk, M. I., & Bosker, R. J. (2016). Differentiation practices in grade 2 and 3: Variations in teacher behavior in mathematics and reading comprehension lessons. *The Journal of Classroom Interaction, 51*(2), 50-72.

Rohrbeck, C. A., Ginsburg-Block, M. D., Fantuzzo, J. W., & Miller, T. R. (2003). Peer-assisted learning interventions with elementary school students: A meta-analytic

review. *Journal of Educational Psychology, 95*(2), 240-257. https://doi.org/10.1037/0022-0663.95.2.240

Royer, D. J., Lane, K. L., Dunlap, K. D., & Ennis, R. P. (2019). A systematic review of teacher-delivered behavior-specific praise on k–12 student performance. *Remedial and Special Education, 40,* 112-128. https://doi.org/10.1177/0741932517751054

Rubow, C. C., Vollmer, T. R., & Joslyn, P. R. (2018). Effects of the good behavior game on student and teacher behavior in an alternative school. *Journal of Applied Behavior Analysis, 51*(2), 382-392. https://doi.org/10.1002/jaba.455

Savage, C., Lewis, J., & Colless, N. (2011). Essentials for implementation: Six years of school wide positive behaviour support in New Zealand. *New Zealand Journal of Psychology (Online), 40*(1), 29-37.

Sawka, K. D., McCurdy, B. L., & Mannella, M. C. (2002). Strengthening emotional support services: An empirically based model for training teachers of students with behavior disorders. *Journal of Emotional and Behavioral Disorders, 10*(4), 223-232. https://doi.org/10.1177/10634266020100040401

Schaffer, E., Stringfield, S., & Devlin-Scherer, R. (2017). A two-year follow-up of a staff development program designed to change teacher behavior. *Journal of Classroom Interaction, 52*(1), 35-44.

Schmidt, C., & Stichter, J. P. (2012). The use of peer-mediated interventions to promote the generalization of social competence for adolescents with high-functioning autism and asperger's syndrome. *Exceptionality, 20*(2), 94-113. https://doi.org/10.1080/09362835.2012.669303

Sedova, K. (2017). Transforming teacher behaviour to increase student participation in classroom discourse. *Teacher Development, 21*(2), 225-242. https://doi.org/10.1080/13664530.2016.1224775

Sherman, J. C. (2012). *Positive peer reporting and positive peer reporting combined with tootling: A comparison of interventions* [Doctor of Philosophy, University of Southern Mississippi]. Aquila. https://core.ac.uk/download/pdf/301298828.pdf

Sieberer-Nagler, K. (2016). Effective classroom-management & positive teaching. *English Language Teaching, 9*(1), 163-172.

Sinclair, A. C., Gesel, S. A., & Lemons, C. J. (2019). The effects of peer-assisted learning on disruptive behavior and academic engagement. *Journal of Positive Behavior Interventions, 21*(4), 238-248. https://doi.org/10.1177/1098300719851227

Skinner, B. F. (1968). *The technology of teaching*. Appleton-Century-Crofts.

Skinner, C. H., Cashwell, T. H., & Skinner, A. L. (2000). Increasing tootling: The effects of a peer-monitored group contingency program on students' reports of peers' prosocial behaviors. *Psychology in the Schools, 37*, 263-270. https://doi.org/10.1002/(SICI)1520-6807(200005)37:33.0.CO2-C

Skinner, C. H., Neddenriep, C. E., Robinson, S. L., Ervin, R., & Jones, K. (2002). Altering educational environments through positive peer reporting: Prevention and remediation of social problems associated with behavior disorders. *Psychology in the Schools, 39*(2), 191-202. https://doi.org/10.1002/pits.10030

Skinner, E. A., & Belmont, M. J. (1993). Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year. *Journal of Educational Psychology, 85*(4), 571-581. https://doi.org/10.1037/0022-0663.85.4.571

Slavin, R. (1991). Cooperative learning and group contingencies. *Journal of Behavioral Education, 1*(1), 105-115. https://doi.org/10.1007/BF00956756

Slavin, R. E. (1980). Cooperative learning. *Review of Educational Research, 50*(2), 315-342.

Slavin, R. E. (1995). *Coopertaive learning: Theory, research, and practice* (2nd ed.). Allyn & Bacon.

Slavin, R. E., Hurley, E. A., & Chamberlain, A. (2003). Cooperative learning and achievement: Theory and research. In W. M. Reynolds & G. E. Miller (Eds.), *Handbook of psychology: Educational psychology* (pp. 177-198). Wiley.

Soares, D., Harrison, J., Vannest, K., & McClelland, S. (2016). Effect size for token economy use in contemporary classroom settings: A meta-analysis of single-case research. *School Psychology Review, 45*(4), 379-399. https://doi.org/10.17105/SPR45-4.379-399

Spilt, J., Leflot, G., Onghena, P., & Colpin, H. (2016). Use of praise and reprimands as critical ingredients of teacher behavior management: Effects on children's development in the context of a teacher-mediated classroom intervention. *Prevention Science, 17*(6), 732-742. https://doi.org/10.1007/s11121-016-0667-y

Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis, 10*(2), 349-367. https://doi.org/10.1901/jaba.1977.10-349

Sugai, G., Simonsen, B., Freeman, J., & La Salle, T. (2016). Capacity development and multi-tiered systems of support: Guiding principles. *Australasian Journal of Special Education, 40*(2), 80. https://doi.org/10.1017/jse.2016.11

Sutherland, K. S., Wehby, J. H., & Copeland, S. R. (2000). Effect of varying rates of behavior-specific praise on the on-task behavior of students with EBD. *Journal of Emotional and Behavioral Disorders, 8*(1), 2-8. https://doi.org/10.1177/106342660000800101

Thiemann, K. S., & Goldstein, H. (2001). Social stories, written text cues, and video feedback: Effects on social communication of children with autism. *Journal of Applied Behavior Analysis, 34*(4), 425-446. https://doi.org/10.1901/jaba.2001.34-425

Trussell, R. P., Chen, H. J., Lewis, T. J., & Luna, N. E. (2018). Reducing escape-maintained behavior through the application of classroom-wide practices and individually

designed interventions. *Education and Treatment of Children, 41*(4), 507-532. https://doi.org/10.1353/etc.2018.0027

Turco, T. L., & Elliott, S. N. (1986). Students' acceptability ratings of interventions for classroom misbehaviors: A study of well-behaving and misbehaving youth. *Journal of Psychoeducational Assessment, 4*(4), 281-289. https://doi.org/10.1177/073428298600400404

Van Petegem, K., Aelterman, A., Van Keer, H., & Rosseel, Y. (2008). The influence of student characteristics and interpersonal teacher behaviour in the classroom on student's wellbeing. *An International and Interdisciplinary Journal for Quality-of-Life Measurement, 85*(2), 279-291. https://doi.org/10.1007/s11205-007-9093-7

Van Ryzin, M. J., & Roseth, C. J. (2018). Cooperative learning in middle school: A means to improve peer relations and reduce victimization, bullying, and related outcomes. *Journal of Educational Psychology, 110*(8), 1192-1201. https://doi.org/10.1037/edu0000265

Vannest, K. J., & Ninci, J. (2015). Evaluating intervention effects in single-case research designs. *Journal of Counseling and Development, 93*(4), 403-411. https://doi.org/10.1002/jcad.12038

Veldman, M. A., Doolaard, S., Bosker, R. J., & Snijders, T. A. B. (2020). Young children working together. Cooperative learning effects on group work of children in Grade 1 of primary education. *Learning and Instruction, 67*, 1-13. https://doi.org/10.1016/j.learninstruc.2020.101308

Von Brock, M. B., & Elliott, S. N. (1987). Influence of treatment effectiveness information on the acceptability of classroom interventions. *Journal of School Psychology, 25*(2), 131-144. https://doi.org/10.1016/0022-4405(87)90022-7

Wahl, E., Hawkins, R. O., Haydon, T., Marsicano, R., & Morrison, J. Q. (2016). Comparing versions of the good behavior game: Can a positive spin enhance effectiveness? *Behavior Modification, 40*(4), 493-517. https://doi.org/10.1177/0145445516644220

Walker, H. M., Ramsey, E., & Gresham, F. M. (2003). Heading off disruptive behavior: How early intervention can reduce defiant behavior and win back teaching time. *American Educator, 27*(4), 1-21.

Weyns, T., Verschueren, K., Leflot, G., Onghena, P., Wouters, S., & Colpin, H. (2017). The role of teacher behavior in children's relational aggression development: A five-wave longitudinal study. *Journal of School Psychology, 64*, 17-27. https://doi.org/10.1016/j.jsp.2017.04.008

Wills, H. P., Iwaszuk, W. M., Kamps, D., & Shumate, E. (2014). CW-FIT: Group contingency effects across the day. *Education and Treatment of Children, 37*(2), 191-210. https://doi.org/10.1353/etc.2014.0016

Wirth, O., Slaven, J., & Taylor, M. A. (2014). Interval sampling methods and measurement error: A computer simulation. *Journal of Applied Behavior Analysis, 47*(1), 83-100. https://doi.org/10.1002/jaba.93

Wright, R. A., & McCurdy, B. L. (2012). Class-wide positive behavior support and group contingencies: Examining a positive variation of the good behavior game. *Journal of Positive Behavior Interventions, 14*(3), 173-180. https://doi.org/10.1177/1098300711421008

Zhu, M., Urhahne, D., & Rubie-Davies, C. M. (2018). The longitudinal effects of teacher judgement and different teacher treatment on students' academic outcomes. *Educational Psychology, 38*(5), 648-668. https://doi.org/10.1080/01443410.2017.1412399

Appendices

Appendix A - Informed Consent Form - Teachers

Please retain a copy of this form for your personal records.

Research Project: Improving classroom behaviour within typical New Zealand primary school classrooms, using a tootling intervention.

Name of participant: _____

I have received a copy of the Information Sheet describing the research project and have been given sufficient time to read it. Any questions that I have, relating to the research, have been answered to my satisfaction. I understand that I can ask further questions about the research at any time during my participation, and that I can withdraw my participation at any time (up to two weeks) after completion of data collection.
I understand that I can ask to have the observations stopped at any time.
When I sign this consent form, I will retain ownership of the collected data, but I give consent for the researcher to use the data for the purposes of the research outlined in the Information Sheet.
I understand that my identity will remain confidential in the presentation of the research findings.

Please complete the following checklist. Tick [✓] the appropriate box for each point.	**YES**	**NO**
I have the right to decline to participate in any part of the research activity.		
I know who to contact if I have any questions about the study in general.		
I understand that the information supplied by me could be used in future academic publications.		
I consent to being interviewed at the beginning of the research regarding what I value within my classroom and for my students.		
I consent to completing a questionnaire near the end of the study, as a post-intervention measure on my thoughts about the intervention.		
I consent to having up to two trained observers in my classroom during times agreed upon by myself and the lead researcher.		
I wish to receive a copy of the findings		

Participant:	_____	Researcher:	_____
Signature:	_____	Signature:	_____
Date:	_____	Date:	_____
Contact Details:	_____	Contact Details:	_____
	_____		_____

Appendix B - Informed Consent - Individual Student

Consent Form

Please retain a copy of this form for your personal records.

Research Project: Improving classroom behaviour within typical New Zealand primary school classrooms, using a tootling intervention.

Name of participant: _____

I have received a copy of the Information Sheet describing the research project and have been given sufficient time to read it. Any questions that I have, relating to the research, have been answered to my satisfaction.
When I sign this consent form, I will retain ownership of the collected data, but I give consent for the researcher to use the data for the purposes of the research outlined in the Information Sheet.
I understand that my identity and the identity of my child will remain confidential in the presentation of the research findings.

Please complete the following checklist. Tick [✓] the appropriate box for each point.	**YES**	**NO**
I have the right to decline to have my child's results reported at any level.		
I know who to contact if I have any questions about the study in general.		
I understand that the information supplied by me could be used in future academic publications.		
I understand that participation in this study is confidential and that no material, which could identify me or my child personally, will be used in any reports on this study.		
I consent to having my child's results reported, anonymously, for the purposes of the research outlined in the Information Sheet.		
I wish to receive a summary of the findings		

Participant:	_____	Researcher:	_____
Parent/Caregiver:	_____	Signature:	_____
Signature:	_____	Date:	_____
Date:	_____	Contact Details:	_____
Contact Details:	_____		_____
	_____		_____
	_____		

Appendix C - Passive Consent Form

School of Psychology

THE UNIVERSITY OF WAIKATO
Te Whare Wānanga o Waikato

Associate Professor Angelika Anderson
Faculty of Social Science
Waikato University
Phone: 07 838 4466 ext 9209
Email: angelika.anderson@waikato.ac.nz

Cheniel Powell
Phone: 0210335182
Email: ceh19@students.waikato.ac.nz

Research Project:
Improving classroom behaviour within typical New Zealand primary school classrooms, using a tootling intervention.

I have read and understood the Information Sheet regarding the above research project and do **NOT** give consent for my child to participate in the research on classroom behaviour.

Participant's name (please print): _____

Signature: _____ Date: _____

Appendix D - On-Task Behaviour Data Collection Form

Date: **Class:** **Teacher:**
Observer: **Time:**

Field notes:

Definition of on-task
Student is; complying with teacher instructions and classroom rules; attending to (eyes oriented towards) appropriate materials, teacher, or peers, or has eyes closed if appropriate to task, such as mindfulness; and is participating in independent seatwork, peer- or group activities, at a reasonable noise level. Independent seat work is defined as 1 or 2 knees or buttocks touching seat, with all four chair legs on the ground, using appropriate stationary or equipment in an appropriate manner. For example, if task is writing with a pencil, student is holding pencil using pencil grip, while attending to work book, or with pencil touching paper. If working on a laptop, fingers must be positioned touching keys or approximately 5cm above keys.
Examples:

1. Sitting quietly on the mat, looking at the teacher, and answering questions when prompted.
2. Sitting on chair, at desk, with pen in hand, looking at academic materials or writing in academic book.
3. Sitting on chair, at desk, nothing in hands, looking at teacher or white board.
4. Talking one-to-one with teacher.
5. Talking to a peer or peers if granted permission by teacher.
   (All behaviour must be in accordance with teacher instruction to be counted as on-task.

Non-examples:

1. Sitting quietly on mat, gazing in a direction other than the teacher, white board, or other material instructed by the teacher.
2. Sitting on chair, at desk, engaging in verbal behaviour after teacher has requested silence.

Student name						
Interval	**1**	**2**	**3**	**4**	**5**	**6**
Student num.	1	2	3	4	5	6
On-task ✓ / ✗						
Notes						

Interval	**7**	**8**	**9**	**10**	**11**	**12**
Student num.	1	2	3	4	5	6
On-task ✓ / ✗						
Notes						

Interval	**13**	**14**	**15**	**16**	**17**	**18**
Student num.	1	2	3	4	5	6
On-task ✓ / ✗						
Notes						

Student name						
Interval	**19**	**20**	**21**	**22**	**23**	**24**
Student num.	1	2	3	4	5	6
On-task ✓ / ✕						
Notes						

Interval	**25**	**26**	**27**	**28**	**29**	**30**
Student num.	1	2	3	4	5	6
On-task ✓ / ✕						
Notes						

Interval	**31**	**32**	**33**	**34**	**35**	**36**
Student num.	1	2	3	4	5	6
On-task ✓ / ✕						
Notes						

Interval	**37**	**38**	**39**	**40**	**41**	**42**
Student num.	1	2	3	4	5	6
On-task ✓ / ✕						
Notes						

Interval	**43**	**44**	**45**	**46**	**47**	**48**
Student num.	1	2	3	4	5	6
On-task ✓ / ✕						
Notes						

Interval	**49**	**50**	**51**	**52**	**53**	**54**
Student num.	1	2	3	4	5	6
On-task ✓ / ✕						
Notes						

Interval	**55**	**56**	**57**	**58**	**59**	**60**
Student num.	1	2	3	4	5	6
On-task ✓ / ✕						
Notes						

Student name						
Interval	**61**	**62**	**63**	**64**	**65**	**66**
Student num.	1	2	3	4	5	6
On-task ✓ / ✕						
Notes						

Interval	**67**	**68**	**69**	**70**	**71**	**72**
Student num.	1	2	3	4	5	6
On-task ✓ / ✕						
Notes						

Interval	**73**	**74**	**75**	**76**	**77**	**78**
Student num.	1	2	3	4	5	6
On-task ✓ / ✕						
Notes						

Interval	**79**	**80**	**81**	**82**	**83**	**84**
Student num.	1	2	3	4	5	6
On-task ✓ / ✕						
Notes						

Interval	**85**	**86**	**87**	**88**	**89**	**90**
Student num.	1	2	3	4	5	6
On-task ✓ / ✕						
Notes						

Interval	**91**	**92**	**93**	**94**	**95**	**96**
Student num.	1	2	3	4	5	6
On-task ✓ / ✕						
Notes						

Interval	**97**	**98**	**99**	**100**	**101**	**102**
Student num.	1	2	3	4	5	6
On-task ✓ / ✕						
Notes						

Student name						
Interval	**103**	**104**	**105**	**106**	**107**	**108**
Student num.	1	2	3	4	5	6
On-task ✓ / ✕						
Notes						

Interval	**109**	**110**	**111**	**112**	**113**	**114**
Student num.	1	2	3	4	5	6
On-task ✓ / ✕						
Notes						

Interval	**115**	**116**	**117**	**118**	**119**	**120**
Student num.	1	2	3	4	5	6
On-task ✓ / ✕						
Notes						

Interval	**121**	**122**	**123**	**124**	**125**	**126**
Student num.	1	2	3	4	5	6
On-task ✓ / ✕						
Notes						

Interval	**127**	**128**	**129**	**130**	**131**	**132**
Student num.	1	2	3	4	5	6
On-task ✓ / ✕						
Notes						

Interval	**133**	**134**	**135**	**136**	**137**	**138**
Student num.	1	2	3	4	5	6
On-task ✓ / ✕						
Notes						

Interval	**139**	**140**	**141**	**142**	**143**	**144**
Student num.	1	2	3	4	5	6
On-task ✓ / ✕						
Notes						

Total intervals
student on-task:

Total intervals
student observed:

Appendix E - Data Collection Form - Teacher Behaviour

Date: ***Class:*** ***Teacher:***

Observer: ***Time:***

Field notes:

Teacher praise: a positive verbal statement or gesture indicating approval of appropriate behaviour. For example, "Well done", "you're sitting beautifully", or "I love the way you stopped, looked and listened". Non-examples include giving out tokens with no explanation and praise in response to correct academic responses.

Corrective Statements: a verbal statement or gesture which functions to stop or change inappropriate student behaviour. (This function of behaviour can be assumed if teacher behaviour stops when student behaviour stops or changes). Includes requests to stop inappropriate behaviour, such as "Please stop doing that" and "Don't throw darts", as well as threats, such as "Luke, this is your final warning" and "Claire, if you can't settle, you're going to have to move". Non-examples include chanting, "Tahi, rua" or "Shh, shh, sh, sh, sh" to gain the attention of all students.

Group: if two-or-more students are the target of teacher behaviour, but teacher does not vocalise specific names. For example "You guys need to stop" or "Well done to the children who are still on task".

Individual: teacher looks at and/or states the name of the student they are addressing. More than one student can be named at the same time. For example, "Sadie, John, and Fraser, please stop" or "James, thanks for putting your hand up".

Specific: teacher labels a specific behaviour. For example "I love how quietly you are sitting" or "Do not put tokens in your mouth".

General : the teacher does not label the behaviour. For example, "Well done, Jimmy" or "Stop that please".

Interval	Behaviour	Target	Type	Notes
	☐ Praise ☐ Correction	☐ Group ☐ Individual	☐ Specific ☐ General	
	☐ Praise ☐ Correction	☐ Group ☐ Individual	☐ Specific ☐ General	
	☐ Praise ☐ Correction	☐ Group ☐ Individual	☐ Specific ☐ General	
	☐ Praise ☐ Correction	☐ Group ☐ Individual	☐ Specific ☐ General	
	☐ Praise ☐ Correction	☐ Group ☐ Individual	☐ Specific ☐ General	
	☐ Praise ☐ Correction	☐ Group ☐ Individual	☐ Specific ☐ General	
	☐ Praise ☐ Correction	☐ Group ☐ Individual	☐ Specific ☐ General	
	☐ Praise ☐ Correction	☐ Group ☐ Individual	☐ Specific ☐ General	

	☐ Praise ☐ Correction	☐ Group ☐ Individual	☐ Specific ☐ General	
	☐ Praise ☐ Correction	☐ Group ☐ Individual	☐ Specific ☐ General	
	☐ Praise ☐ Correction	☐ Group ☐ Individual	☐ Specific ☐ General	
	☐ Praise ☐ Correction	☐ Group ☐ Individual	☐ Specific ☐ General	
	☐ Praise ☐ Correction	☐ Group ☐ Individual	☐ Specific ☐ General	
	☐ Praise ☐ Correction	☐ Group ☐ Individual	☐ Specific ☐ General	
	☐ Praise ☐ Correction	☐ Group ☐ Individual	☐ Specific ☐ General	
	☐ Praise ☐ Correction	☐ Group ☐ Individual	☐ Specific ☐ General	
	☐ Praise ☐ Correction	☐ Group ☐ Individual	☐ Specific ☐ General	
	☐ Praise ☐ Correction	☐ Group ☐ Individual	☐ Specific ☐ General	
	☐ Praise ☐ Correction	☐ Group ☐ Individual	☐ Specific ☐ General	
	☐ Praise ☐ Correction	☐ Group ☐ Individual	☐ Specific ☐ General	
	☐ Praise ☐ Correction	☐ Group ☐ Individual	☐ Specific ☐ General	
	☐ Praise ☐ Correction	☐ Group ☐ Individual	☐ Specific ☐ General	
	☐ Praise ☐ Correction	☐ Group ☐ Individual	☐ Specific ☐ General	
	☐ Praise ☐ Correction	☐ Group ☐ Individual	☐ Specific ☐ General	
	☐ Praise ☐ Correction	☐ Group ☐ Individual	☐ Specific ☐ General	
	☐ Praise ☐ Correction	☐ Group ☐ Individual	☐ Specific ☐ General	
	☐ Praise ☐ Correction	☐ Group ☐ Individual	☐ Specific ☐ General	
	☐ Praise ☐ Correction	☐ Group ☐ Individual	☐ Specific ☐ General	

Number of: Praise, Group, Specific: Praise, Group, General: Praise Individual, Specific:
Praise, Individual, General: Correction, Group, Specific: Correction, Group, Specific:
Correction, Individual, Specific: Correction, Individual, General:

Appendix F - Procedural Integrity - Teacher Training

Date:_____ Teacher:_____

Tick appropriately once item has been discussed with teacher.

		(✓)
1.	**Introduce tootling**	
	Provide script for teachers	___
	Explain what a "tootle" is	___
2.	**Explain each step of the procedure (from the script)**	
	Step 1 - Introduce tootling	___
	Step 2 - Discuss examples	___
	Step 3 - Give a written example	___
	Step 4 - Explain procedure	___
	Step 5 - Explain where tootle go and when	___
	Step 6 - Explain the team work, tootle target, and class reward	___
	Step 7 - Discuss rewards	___
	Step 8 - Any questions	___
3.	**Rehearse script**	
	Allow teacher to practice script as many times as needed	___
	Provide correctional feedback if/where needed	___
4.	**Any questions?**	
	Answer any questions teacher has	___

Number of items completed /13 Treatment integrity percentage: _____

Date: _____

Observer: _____

Appendix G - Procedural Integrity - Student Training

Date: _____ Teacher: _____ Class: _____

Did the teacher…

	Y/N
1. Discuss examples and non-examples	___
2. Give at least one written example	___
3. Explain the procedure	___
4. Explain where to put tootles and when	___
5. Discuss team work, tootle target, and class reward	___
6. Discuss potential rewards	___
7. Ask for and answer questions	___
8. Allow everyone to practice tootling and provide appropriate feedback	___

Number of items completed: /8 Treatment integrity percentage: _____

Date: _____

Observer: _____

Appendix H - Treatment Integrity - Teacher

Teacher: Class:

Please tick appropriately when each item has been completed daily

DATE:											
Prior to tootling each day:											
1. Tootle cards placed within easy access											
2. Tootle container accessible											
3. Progress chart visible											
4. Tootling procedure reviewed (if needed)											
Once tootling has finished for the day:											
5. Read several tootles and provide praise/corrective feedback											
6. Count tootles (reward if target achieve)											
7. Update progress chart											
Total number of tootles collected today											

Appendix I - Treatment Integrity - Researcher

Date													
Class													
Tootle cards on student desks													
Tootling collection container visible													
Feedback chart updated													
Feedback chart visible to whole classroom													
Tootles counted													
Selection of tootles read aloud													
Reward issued													

Date													
Class													
Tootle cards on student desks													
Tootling collection container visible													
Feedback chart updated													
Feedback chart visible to whole classroom													
Tootles counted													
Selection of tootles read aloud													
Reward issued													

Appendix J - Teacher Script

Please read the following script to the students, which will explain how tootling will work in your classroom.

We're going to introduce the tootling challenge into our classroom to help us be kind to each other and record when our classmates do good things. I'll give you all some note cards and when you see another student in the class doing something good you can write it down on one of the cards, place your card in this container *(show container in the position it will be during the procedure)*, then add a notch on this progress chart *(show progress chart, location during tootling, and how to move mark up after each tootle)*. This is called tootling. When you write your tootle on the card, it is important that you write your name in the top corner, then write the name of the student who did something good and what it was that they did. Here is an example *(show example).* Tootles do not need to be a secret so you can tell people when you are writing about them, so long as this does not disrupt the class.

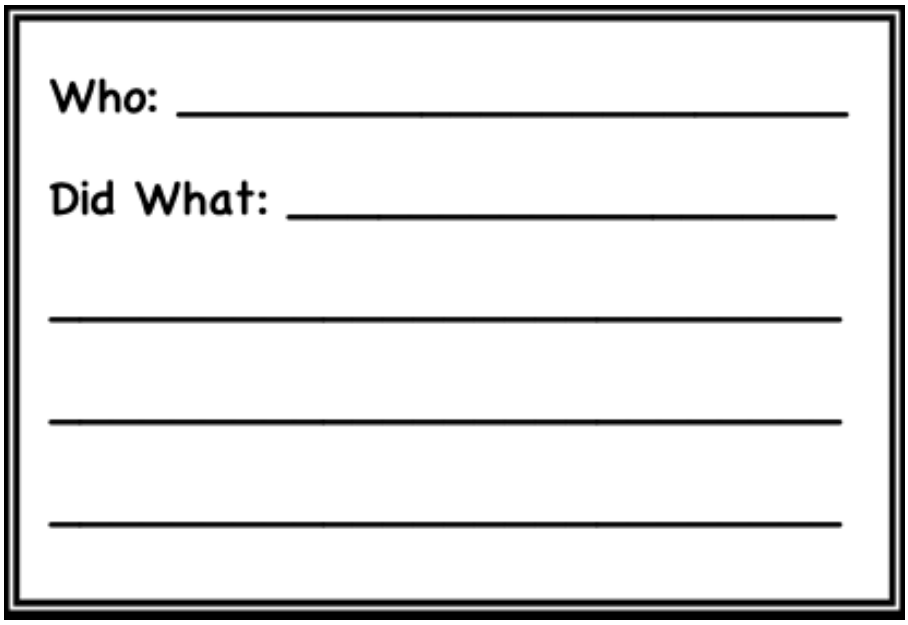
Now, here's the challenge - everyone in the class must work together as a team to try to reach a target number of tootles. If the class reaches the target you will all receive a reward. If the target is not reached, all that day's tootles will be carried over to the next day's target. The target will be different each day and will be randomly chosen at the start of the lesson but will be kept a secret until the end of the lesson. So, it's important that you are all well-behaved in class and try to write as many tootles as you can to help your team. At the end of each lesson I will quietly read some of your tootles and let you know the sort of awesome behaviour you've noticed about each other.

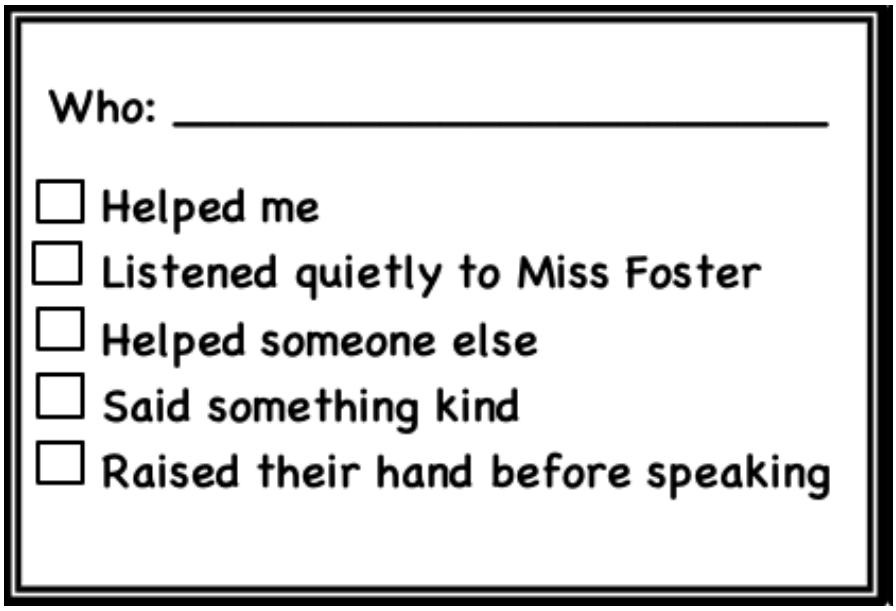
So, let's talk about the rewards! I have some ideas for rewards which I think you might enjoy, but first I'd like to hear if you have any ideas. Remember, they need to be appropriate for the whole class. *(Let the children give you their suggestions. If any of these matches with your own ideas or are good ideas that you haven't already thought of, agree to these rewards. There can be a different reward each time or you can keep to the same reward. If you don't like any of the children's suggestions, make your own to see if they can agree on any of yours. It is important that you AND the children are happy with the rewards).*
Now we will talk about the sorts of things you can and can't write down about other students *(this can either be done as a class discussion or you can give 3-4 examples of correct and incorrect tootles).*

Any questions?

Finally, I'd like you all to practice writing one tootle on a piece of paper and posting it in the tootle container. *(Read each student's example and provide feedback on whether each is correct. Read several correct examples out loud to the class).*

Appendix K - Sample Tootle Cards





Appendix L - Tootle Chart

Our Tootle Progress Chart

91	92	93	94	95	96	97	98	99	100
81	82	83	84	85	86	87	88	89	90
71	72	73	74	75	76	77	78	79	80
61	62	63	64	65	66	67	68	69	70
51	52	53	54	55	56	57	58	59	60
41	42	43	44	45	46	47	48	49	50
31	32	33	34	35	36	37	38	39	40
21	22	23	24	25	26	27	28	29	30
11	12	13	14	15	16	17	18	19	20

Appendix M - Modified BIRS

Please evaluate the tootling intervention by circling the number which best describes your agreement or disagreement with each statement. Please answer every question.

		Strongly disagree	Disagree	Slightly disagree	Slightly agree	Agree	Strongly agree
1.	Tootling was an acceptable intervention for children's problem behaviour.	1	2	3	4	5	6
2.	Most teachers would find tootling appropriate for a variety of behaviour problems.	1	2	3	4	5	6
3.	Tootling proved effective in changing children's problem behaviour.	1	2	3	4	5	6
4.	I would suggest the use of tootling to other teachers.	1	2	3	4	5	6
5.	Behaviour in the classroom was severe enough to warrant the use of tootling.	1	2	3	4	5	6
6.	Most teachers would find tootling suitable for improving general classroom behaviour.	1	2	3	4	5	6
7.	I would be willing to use tootling in the classroom again.	1	2	3	4	5	6
8.	Tootling resulted in negative side-effects for some children.	1	2	3	4	5	6
9.	Tootling was appropriate for a variety of children.	1	2	3	4	5	6
10.	Tootling was consistent with other strategies I have used in the classroom setting.	1	2	3	4	5	6
11.	Tootling was a fair way to handle children's problem behaviour.	1	2	3	4	5	6
12.	Tootling was reasonable for the behaviour problems experienced in my classroom.	1	2	3	4	5	6
13.	I liked the procedures used in tootling.	1	2	3	4	5	6
14.	Tootling was a good way to handle classroom behaviour.	1	2	3	4	5	6
15.	Overall, tootling was beneficial for the children in my classroom.	1	2	3	4	5	6
16.	Tootling quickly improved the children's behaviour.	1	2	3	4	5	6
17.	Tootling will produce a lasting improvement in the children's behaviour.	1	2	3	4	5	6
18.	Tootling improved the children's behaviour to the point that it was not noticeably deviate from other children's behaviour.	1	2	3	4	5	6
19.	Soon after using tootling, I noticed a positive change in problem behaviour.	1	2	3	4	5	6
20.	The children's behaviour will likely remain at an improved level even after tootling is discontinued.	1	2	3	4	5	6
21.	Using tootling not only improved the children's behaviour in the classroom, but also in other settings (e.g., other classrooms, home).	1	2	3	4	5	6
22.	When comparing the children in my classroom with well-behaved peers before and after use of tootling, the children's and the peers' behaviour was more alike after using tootling.	1	2	3	4	5	6
23.	Tootling produced enough improvement in the children's behaviour that behaviour is no longer a problem in the classroom.	1	2	3	4	5	6
24.	Other behaviours related to the problem behaviour also improved as a result of tootling.	1	2	3	4	5	6

Appendix N - Modified CIRP

What do you think about tootling?

I'd love to know what you think about tootling! Please circle the smiley face which shows how much you agree or disagree with each sentence below. A red face tells me you do **not** agree with the sentence. A green face tells me you **do** agree with the sentence and the other faces are all somewhere in-between.

		I do not agree					I agree
1	Tootling is fair						
2	Tootling was too hard on me						
3	Tootling caused problems with my friends						
4	There are better ways to handle problem behaviour than tootling						
5	Tootling would help other children too						
6	I liked tootling						
7	I think tootling would help me do better in school						